

| | 201 | 211 | 221 | 231 | 241 | 251 | 261 | 271 | 281 | 291 |
|------------|-------------|------------|------------|-------------|------------|-------------|------------|------------|-----------|------------|
| SEQ 3 | LAHAGRKATT | VAPW | ----- | ISPS | AIATEKVGCG | PDRVKGPGDI | P----- | ----- | ----- | ----- |
| SEQ 6 | IAHAGRKAST | VAPW | ----- | LSAN | DTASEKMGCH | PDIVKGVPTN | P----- | ----- | ----- | ----- |
| SEQ 8 | LAHAGRKASA | VAPW | ----- | LAAQAGKSS | LKADSEVGGW | PADVVGPSGG | E----- | ----- | ----- | ----- |
| SEQ 10 | IAHAGRKASIN | TAPW | ----- | LMNKG | IVATEKVGCG | PDVIGPSTV | P----- | ----- | ----- | ----- |
| SEQ 12 | IGHGGRKASG | QPLF | ----- | LHLE | QVADKSVNGF | ADKAVAPSAL | A----- | ----- | ----- | ----- |
| SEQ 14 | LAHAGRKAST | LAPW | ----- | ---QIQHGW | QEHCVGPSTE | P----- | ----- | ----- | ----- | ----- |
| SEQ 16 | LAHAGRKAST | KAPW | ----- | HYQRGKS | ELAGPEQGGW | PENVWAPSAL | S----- | ----- | ----- | ----- |
| SEQ 19 | INHPRQSPM | GAGT | ----- | ----- | RGLW | E-KAVAPSVP | P----- | ----- | ----- | ----- |
| SEQ 22 | LAHAGRKAST | KAPWHDSPPT | SGEYKPRBGL | QVVGPEYGGW | PDDVWAPSAL | P----- | ----- | ----- | ----- | ----- |
| SEQ 24 | LAHAGRKAST | TAPY | ----- | RG-Y | TVATEAQGGW | ENDVVGPFPTN | E----- | ----- | ----- | ----- |
| SEQ 27 | LAHAGRKAST | LAPW | ----- | ITEARGK | ALAQSEMGW | PDDVWAPSAL | P----- | ----- | ----- | ----- |
| SEQ 30 | LSHAGRKASC | VSPW | ----- | LSVN | AVAAEEVGGW | PDNIWAPSAL | A----- | ----- | ----- | ----- |
| SEQ 33 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 35 | IGHAGRKAST | VVPW | ----- | LDRK | NTAF? | ----- | ----- | ----- | ----- | ----- |
| SEQ 38 | LSHAGRKASC | VSPW | ----- | LSIN | AVAAKEVGGW | PDNIWAPSAL | A----- | ----- | ----- | ----- |
| SEQ 40 | LAHAGRKAST | VAPW | ----- | LSGG | DVAGEDVNGW | PDDVWAPSAL | P----- | ----- | ----- | ----- |
| SEQ 42 | ----- | ----- | ----- | ----- | LPS | KRAGKEAGGW | PEDVVGPSGG | EDPTWDERSS | ----- | ----- |
| SEQ 44 | VGHPRQARG | SVO | ----- | QHPISASD | VOLKQEM | ----- | ----- | ----- | ----- | ----- |
| SEQ 83 | IGHAGRKASC | VAPW | ----- | LDAG | LAAEKAAAGW | PDDVVGPSNE | P----- | ----- | ----- | ----- |
| SEQ 85 | LAHAGRKASD | WSPF | ----- | YRGEKKQ | KFTVQEDGGW | PDVWAPSAL | A----- | ----- | ----- | ----- |
| Bacteria | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| T44612 | IAHAGRKASA | NRPW | ----- | EGDD | HIGADDARGW | ETIAPSAI | A----- | ----- | ----- | ----- |
| NP_625402 | LAHAGRKAST | AQPW | ----- | RGG | APVGADAYGW | QPLAPSAI | A----- | ----- | ----- | ----- |
| NP_295913 | LAHAGRKAST | YAPW | ----- | RGK | GAVPAELGGW | QVIGPDEN | S----- | ----- | ----- | ----- |
| AF320254 | IGHAGRKAT | KLAW | ----- | EG | IDEPLEAGAW | ELISASPL | P----- | ----- | ----- | ----- |
| OYE family | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| AF4875 | LWATGRAADP | DVLA | ----- | DMK-D | LISSS-AVPV | SEKGP | ----- | ----- | ----- | ----- |
| AF4961 | LMHAGRATIP | QMTG | ----- | SPAVSAS | ATVWDSPTCC | YSHIP | ----- | ----- | ----- | ----- |
| C62460 | LMYLGRVANP | KDLK | ----- | DAGLPL | IGPSA-VYW | DEESE | ----- | ----- | ----- | ----- |
| Nc4452 | LMHSLGRAANP | DVLA | ----- | KBGLK | LKSSS-AVPV | KEGAP | ----- | ----- | ----- | ----- |
| ScOYE1 | LMVLGWAAPF | DTLA | ----- | RDG-LR | YDSASDNVYM | NABQ | ----- | ----- | ----- | ----- |
| ScOYE2 | LMVLGWAAPF | DTLA | ----- | RDG-LR | YDSASDNVYM | NABQ | ----- | ----- | ----- | ----- |
| ScOYE3 | LMVLGWAAPF | DVLA | ----- | RDG-LR | YDCASDRVYM | NATLQ | ----- | ----- | ----- | ----- |
| A36990 | LMYLGRVANA | KDLK | ----- | DSG-LP | LIAPS-AVPV | DENSE | ----- | ----- | ----- | ----- |
| SEQ 3 | TKRAIAA-GA | DFVEIHNAHG | YLLSSFLSP | -AANNRTDOY | G-GSPENRIR | LSLEIAQLTR | DAVGPHVP | ----- | VPLR | ISAS-DWCE |
| SEQ 6 | VKRAVKA-GA | DFVEIHNAHG | YLLMSFLSP | -AVNTRTDEY | G-GSPENRIR | LSLEIAKLTR | ENVPKDMP | ----- | VPLR | VSAT-DWLE |
| SEQ 8 | ARLAVQA-GV | DFVEIHNAHG | YLINEFLSP | -VTNKRTRDAY | G-GSPENRIR | IVREVAIAIR | AVIEGPM | ----- | LPLR | ISAT-EWLE |
| SEQ 10 | CKRAIAA-GA | DFVEIHNAHG | YLLSSFLSP | -SSNTRTDEY | G-GSPENRIR | LSLEIAQVTR | DAVGPNVP | ----- | VPLR | VSAT-DWIE |
| SEQ 12 | ARRAVEISGF | DAVEIHGAHG | YLINEFPYSP | -ISNKRTRDEY | G-GSPENRIR | FLKEVIDSVK | SSIPNDVP | ----- | VPLR | ISAA-ENSP |
| SEQ 14 | AMRAVEISKF | DAVEIHGAHG | CLIHQFLSK | -LTNKRADQY | G-GSPENRIR | FLKQIENIK | RKIET | ----- | IFLK | FPMS-DNCS |
| SEQ 16 | AQRAKKA-GF | DLISITHAAG | YLISEFLSP | -ISNKRTRDAY | G-GSPENRIR | VLRBIISAVR | SVIPEMDP | ----- | LPVR | VSAT-EMME |
| SEQ 19 | ARITABA-GF | NGVEIHNAHG | YLLAQFLSK | -KTNRRGDEY | G-GSLENRRR | IVGELIKR | ROVTEAGVEE | BAKPKVGIK | LMSA-DWQA | -GRDKKEEE |
| SEQ 22 | AKRAIBA-GV | DFVEIHNAHG | YLITEFLSP | -LSNKRTRDAY | G-GSPENRIR | VLIDIIKAVR | AVIPEEM | ----- | PLVR | ISAT-EMME |
| SEQ 24 | AKRAVEI-GF | DFVEIHGAHG | YLISSTVSPA | PTTNDNRNDY | G-GTEKRIL | FPMEVHSVR | KAIPDSMP | ----- | LPVR | VTAT-DWLP |
| SEQ 27 | AKRSNRA-GF | DFVEIHNAHG | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 30 | AKRAIHA-GF | DFVEIHNAHG | YLLHQFLSP | -VSNQRTDEY | G-GSPENRIR | VVLEILDILR | AAIPETTP | ----- | VLVR | VSAT-DWFEP |
| SEQ 33 | ----- | ----- | ----- | TDEY | G-GSPENRIR | VVLEILDILR | AAIPETTP | ----- | ----- | ----- |
| SEQ 35 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 38 | KRA-RA-GF | DFVEIHNAHG | Y-LHQFLSP | -VSNQRTDEY | G-GSPENRIR | VVLEII | ----- | ----- | ----- | ----- |
| SEQ 40 | VKRAKKA-GF | DFVEIHNAHG | YLLHEPICL | -RATPGTST | G-GSPENRIR | LTMSRRPCP | OHV | ----- | ----- | ----- |
| SEQ 42 | AKRAVKA-GV | DFVEIHNAHG | YLLHEFLSP | -ITNKRTRDAY | G-GSPENRIR | LLLEIVIAVR | AAMPSSMP | ----- | LPVR | VSAT-EMME |
| SEQ 44 | AEYLKKA-GF | DFVEIHNAHG | YLLAQFLSK | -TTNKRTRDAY | G-GSLENRRR | LLLEIVIAVR | RTSKRNF | ----- | ILGK | INSV-EFQE |
| SEQ 83 | VKRAVEA-GF | DTIDPHPAHG | YLVSSFLSP | -ATNKRTRDAY | G-GSPENRIR | LALIVIAVR | AVMPEDMP | ----- | LPTR | ISAT-DWLE |
| SEQ 85 | ARNAFEA-GY | DYVEIHNAHG | YLMHSFLSP | -LTNKRTRDAY | G-GSLENRRR | FLLEIVIAVR | QEPFNKG | ----- | LMVR | VSAT-DWAD |
| Bacteria | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| T44612 | ARRARDA-GF | EWELHFAHG | YLGQSFFSE | -HSNKRTRDAY | G-GSPDNRRR | FLLEIVIAVR | EVMPENLP | ----- | LTAR | FGVL-EYDG |
| NP_625402 | ARRALAA-GF | EIAEIHGAHG | YLIHEFLSP | -HSNKRTRDAY | G-GSYANRRR | PALEVVDAVR | EVMPDDKP | ----- | LPFR | VSAT-DWLE |
| NP_295913 | ARRAQVA-GF | DAVEIHNAHG | YLLHQFLSP | -LANTRTDDY | G-GSPENRIR | LLLEIVIAVR | HYWPAHL | ----- | LPVR | VSAT-DWAE |
| AF320254 | TRMAABA-GF | DILELHAHG | YLLSSFLSP | -LTNKRTRDAY | G-GSLENRRR | FLLEIVIAVR | AMMPTNRP | ----- | MSVR | LSCH-DWFP |
| OYE family | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| AF4875 | AKNAINA-GF | DGVEIHGAHG | YLIHQPTOK | -SCNHRQDNR | G-GSLENRRR | FAVEVTRAVI | EAVGADR | ----- | VGVR | LSPY-SOYL |
| AF4961 | AKTAMEI-GF | DGVEIHGAHG | YLPDPLSS | -NVNKRTRDAY | G-GSPKRCR | FVLEIMDELA | ATVGEDN | ----- | LAIR | LSPP-GLFN |
| C62460 | AKRAIBA-GF | DYVEIHGAHG | YFLDPLNP | -ASNKRTRDAY | G-GSLENRRR | LLLRIDKLI | GIVGAEK | ----- | LAIR | LAPW-SSFL |
| Nc4452 | AKNAVEA-GF | DGVEIHGAHG | YLIHQFLDQ | -TCNKRTRDAY | G-GSLENRRR | FAVEVVKAVV | EAVGAEK | ----- | TGIR | LSPP-STFP |
| ScOYE1 | AKNSIAA-GA | DGVEIHGAHG | YLLNQFLDQ | -HSNKRTRDAY | G-GSLENRRR | FTLEVVDALV | BAIGHK | ----- | VGIR | LSPP-GVFN |
| ScOYE2 | AKNSIAA-GA | DGVEIHGAHG | YLLNQFLDQ | -HSNKRTRDAY | G-GSLENRRR | FTLEVVDALV | DAIGHK | ----- | VGIR | LSPP-GVFN |
| ScOYE3 | AKNSIAA-GA | DGVEIHGAHG | YLLNQFLDQ | -HSNKRTRDAY | G-GSLENRRR | FTLEVVDALV | ETIGPER | ----- | VGIR | LSPP-GTFN |
| A36990 | AKHALEA-GF | DYVEIHGAHG | YLLDQFLNL | -ASNKRTRDAY | G-GSLENRRR | LLLRVVDKLI | EVVGANR | ----- | LALR | LSPP-ASQP |

| | 401 | 411 | 421 | 431 | 441 | 451 | 461 | 471 | 481 | 491 |
|------------|------------|------------|------------|------------|------------|-----------|-------------|------------|------------|-----------|
| | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 3 | SWKSEDTVR | PAQELVVK-Q | GAVDLIDISS | GGVLAQO | ----- | KI | KSGPAPQVPP | AVAVKKAAGD | ----- | KLLVAAV |
| SEQ 6 | SWRGVDTVR | FAKILA-RT | GYVDVLDVSS | GGTHSEQ | ----- | HI | HAKPGQAPPP | ALAVKNAAGD | ----- | KLAVASV |
| SEQ 8 | SWDM-QSSL | ELVKRLP-E | WGIDLDVSS | AAHKKDQ | ----- | HI | NHPTAYQTDL | AGQIRAT | ----- | RAAGAST |
| SEQ 10 | SWKLSDSVR | FAELAA-Q | GAIDLDVSS | GGVHAAQ | ----- | KI | KSGPAPQAPPP | AVAIKKAAGD | ----- | KLLVATV |
| SEQ 12 | AWTIEDSKK | LADILV-E | KGIALVDVSS | GGNDYRQPP | ----- | RSGISK | ELREPIHVPL | SRAIKQHVGD | ----- | KLLVSCV |
| SEQ 14 | AWSTEDALK | LADILV-E | LGVKVLDVTS | GGNVAHCKS | ----- | RYLLND | DKQLPSQVPL | ARKLKSHIRN | ----- | RCLLIACS |
| SEQ 16 | SWDLQQTII | ELAKILP-D | LGVDLLDVSS | GGNNKDD | ----- | KI | NVHTYYQIDM | ABQIRAAVHE | AGKQLLVGAV | GLVT |
| SEQ 19 | TDTAREVLK | QIELFE-Q | WGIDFVEVSG | GSYEDPQMAN | GPKEKSERT | MAEAFPLEF | AKIIRTK | ----- | FPKLPLMVT | GGFR |
| SEQ 22 | SWDLEQSTQ | LAKLLP-D | LGVDLLDVSS | GGNSVAQ | ----- | KI | ELTPYYQIDL | AAKIREAVGD | ----- | RLLIQAV |
| SEQ 24 | GWEIEDTVAP | TLAARLR-D | GGVDLDVSS | GGNNKDD | ----- | RI | EVKDCYQVPP | AEKIKDQVNG | ----- | ILIGAV |
| SEQ 27 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 30 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 33 | SWTVQETC- | QLARILP-K | HGVLDVSS | GQIHPKS | ----- | AI | KSGPAYQVDL | AKQVKAAGD | ----- | SVLVS |
| SEQ 35 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 38 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 40 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 42 | SWDVSTIK- | ISKILA-D | LGVDLLDVSS | GGNNHPO | ----- | KI | NMFNT | ----- | ----- | ----- |
| SEQ 44 | FKP-EBAVQ | LCAEALAGM | ----- | DPVETSG | GTYESFG | ----- | FAHRKES | RKRENYPIEF | ABVIRKAVKH | ----- |
| SEQ 46 | TWTLQESIK | LAHQLA-D | RGVDVLDVSS | GQIHKKM | ----- | KV | AAGPGYQAPL | AAKIKKSVGD | ----- | KMLISTV |
| SEQ 83 | SWTVQETVE | LAKMLQE- | ARVDLDVSS | GGLVPPQ | ----- | KI | TVGAGYQVPP | AKAVRDALAK | ----- | IBPDASKR |
| SEQ 85 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| Bacteria | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| T44612 | EQTLESII | ELARLRFK-A | GGLDLSVSV | GFTIIPET | ----- | NI | PMGPAIMGPI | AERVREAKL | ----- | PVTSAM |
| NP_625402 | GMPDPTVTR | FARDLR-A | HGIDLLDVST | GGNVPRV | ----- | RI | PTGPGYQVPP | AARVKAAGT | ----- | LPVA </td |
| NP_295913 | GMDLEQTVQ | LSKLLK-Y | EGVDVLDVSS | GGLTAAQ | ----- | QI | EVGPGYQVPP | AAVSRARET | ----- | ISVMAV |
| AF320254 | GNTADDAVA | IARLRFK-E | AGADIDCSS | GQVWKG | ----- | QP | VYGRMYQVPP | ADIRINEVGI | ----- | PTLAVG |
| OYE family | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| AF4875 | EL-VQPEY | LIA-OM | RRLDVAYLHL | ANSRWL | ----- | DE | EKPHDPNHE | VFVRVWG-Q | ----- | SS-PILLA |
| AF4961 | EQR-VETWF | LCESLKAHP | ----- | NLSYVSP | IEPRYE | ----- | QIFSYEBK | NPLRSWG | ----- | LSVDLSSFR |
| Ca2460 | EE-THSY | LQQLQORAD | NGQLAYVSL | IEPRVIG | ----- | IFDASL | EDQGRSNEF | AYKIWK | ----- | NFRA |
| Nc4452 | DLIP-OPED | VIRKIN | ----- | GFLAYLHL | TQSRVAGN | ----- | MDVOP | EDDEE-NLAP | AAKLMDG | ----- |
| ScOYE1 | ETGLVAQVAY | VAGELEKRAK | AGKRLAPVHL | VEPRVTNP | ----- | PLTEGE | GEYEGSNDP | VYSIWK | ----- | PVIRA |
| ScOYE2 | ETGLVAQVAY | VAGELEKRAK | AGKRLAPVHL | VEPRVTNP | ----- | PLTEGE | GEYEGSNDP | AYSIMWK | ----- | PIIRA |
| ScOYE3 | EPGLIAQVSY | VAGELEKRAK | AGKRLAPVHL | VEPRVTNP | ----- | SLVEGE | GEYEGSNDP | AYSIMWK | ----- | PIIRA |
| A36990 | EE-THSY | LQQLQORAD | NGQLAYVSL | IEPRVIG | ----- | YDVSL | KDQGRSNEF | AYKIWK | ----- | NPIRA |
| | 501 | 511 | 521 | 531 | 541 | 551 | 561 | 571 | 581 | 591 |
| | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 3 | NGKQ-AN | QILEEQD | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 6 | SAHLANS | LLSEKQ | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 8 | ADEATAAAM | LSGPEPK | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 10 | NGKQ-AN | KLLEEBG | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 12 | RDPELLN | KYLEBGT | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 14 | RDIFKLD | EPIANGD | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 16 | E-DGRVTIQR | ENGAKTR | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 19 | TRQOME | AALSDDD | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 22 | TADI-AR | DVDEQGAEK | VAEAKQTHDT | IEVVSSEHGG | RTKADLVLIA | QPLREPEFV | LRTAHNLGV | ----- | ----- | ----- |
| SEQ 24 | DGLFTAN | EILESCK | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 27 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 30 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 33 | TGHL-AE | EVLQSG | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 35 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 38 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 40 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 42 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 44 | VGAM-VDA | LQGVGG | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 83 | IGTL-AE | EIIAGQ | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 85 | VGMM-BQ | SYDSFNG | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| Bacteria | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| T44612 | PQAE | AALQANO | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| NP_625402 | EPQAE | KILANG | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| NP_295913 | TGA-QAE | AIIQAGD | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| AF320254 | AD-HAN | SIIAAGR | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| OYE family | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| AF4875 | AASAEVTEQ | MMATYTP | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| AF4961 | AGQWQSNW | GVLEBGR | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| Ca2460 | YDAPEKTL | LHLDND | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| Nc4452 | PETAK-HLV | DREFFEK | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| ScOYE1 | LHP-EV | REEVKDK | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| ScOYE2 | LHP-EV | REEVKDK | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| ScOYE3 | LHP-EV | REEVKDK | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| A36990 | YDAPEKTL | INDLND | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |

| | 601 | 611 | 621 | 631 |
|------------|-------|-------|-------|-------|
| SEQ 3 | ----- | ----- | ----- | ----- |
| SEQ 6 | ----- | ----- | ----- | ----- |
| SEQ 8 | ----- | ----- | ----- | ----- |
| SEQ 10 | ----- | ----- | ----- | ----- |
| SEQ 12 | ----- | ----- | ----- | ----- |
| SEQ 14 | ----- | ----- | ----- | ----- |
| SEQ 16 | ----- | ----- | ----- | ----- |
| SEQ 19 | ----- | ----- | ----- | ----- |
| SEQ 22 | ----- | ----- | ----- | ----- |
| SEQ 24 | ----- | ----- | ----- | ----- |
| SEQ 27 | ----- | ----- | ----- | ----- |
| SEQ 30 | ----- | ----- | ----- | ----- |
| SEQ 33 | ----- | ----- | ----- | ----- |
| SEQ 35 | ----- | ----- | ----- | ----- |
| SEQ 38 | ----- | ----- | ----- | ----- |
| SEQ 40 | ----- | ----- | ----- | ----- |
| SEQ 42 | ----- | ----- | ----- | ----- |
| SEQ 44 | ----- | ----- | ----- | ----- |
| SEQ 83 | ----- | ----- | ----- | ----- |
| SEQ 85 | ----- | ----- | ----- | ----- |
| Bacteria | ----- | ----- | ----- | ----- |
| T44612 | ----- | ----- | ----- | ----- |
| NP_625402 | ----- | ----- | ----- | ----- |
| NP_295913 | ----- | ----- | ----- | ----- |
| AF320254 | ----- | ----- | ----- | ----- |
| OYE family | ----- | ----- | ----- | ----- |
| Af4875 | ----- | ----- | ----- | ----- |
| Af4961 | ----- | ----- | ----- | ----- |
| Ca2460 | ----- | ----- | ----- | ----- |
| Nc4452 | ----- | ----- | ----- | ----- |
| ScOYE1 | ----- | ----- | ----- | ----- |
| ScOYE2 | ----- | ----- | ----- | ----- |
| ScOYE3 | ----- | ----- | ----- | ----- |
| A36990 | ----- | ----- | ----- | ----- |

Figure 1. A multiple alignment of the 2031 OR amino acid sequence from *A. fumigatus* (SEQ ID No3) along with related 2031 ORs from other fungi and bacteria (see Example 4) and OYEs. Regions 1-11, marked with * or #, refer to amino acids conserved between ORs but not OYEs.

Fungal 2031 ORs are given by the following SEQ ID No.: *A. fumigatus*, SEQ ID Nos. 3, 6 and 8; *A. nidulans*, SEQ ID No. 10; *C. albicans* SEQ ID Nos. 12 and 14; *N. crassa*, SEQ ID Nos. 16 and 19; *M. grisea* SEQ ID Nos. 22 and 44; *S. pombe* SEQ ID No. 24 (NP_595868); *C. trifolii* SEQ ID No. 27; *F. sporotrichioides* SEQ ID Nos. 30, 33 and 35; *F. graminearum* SEQ ID Nos. 38 and 83; *M. graminicola* SEQ ID Nos. 40 and 42; *U. maydis* SEQ ID No 85.

Bacterial ORs resembling 2031 are:

T44612 (*Pseudomonas putida*), SEQ ID No. 86; NP_625402 (*Streptomyces coelicolor*), SEQ ID No. 87; NP_295913 (*Deinococcus radiodurans*), SEQ ID No. 88; AF320254 (*Azoarcus evansii*, SEQ ID No. 89).

Fungal ORs similar to the Old Yellow Enzyme family (originally identified in *S. cerevisiae*):

A. fumigatus, Af4875 and Af4961, SEQ ID Nos. 90 and 91 respectively; *C. albicans*, Ca2460 and A36990, SEQ ID Nos. 92 and 93 respectively; *N. crassa*, Nc4452, SEQ ID No. 94; *S. cerevisiae*, OYE1, OYE2 and OYE3, SEQ ID Nos. 95-97 respectively.

Details of the sequence searches that identified the ORs other than SEQ ID No. 3, and methods for the construction of multiple alignments are given in Example 4 hereinafter.

| | 1 | 11 | 21 | 31 | 41 | 51 | 61 | 71 | 81 | 91 |
|--------|------------|------------|-------------|------------|------------|------------|------------|------------|-------------|-------------|
| SEQ 1 | GTTCGACGTC | ATTGCCACGT | TTGACCCAA | GGGCAGACGC | CATGTCGCGG | AGCGATCGCC | GCGATATGCC | TCGAATTGTC | GCCATTTCGGC | ATCCAGTTTC |
| SEQ 2 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 4 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 5 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 7 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 9 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 11 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 13 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 15 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 17 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 18 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 20 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 21 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 23 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 25 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 26 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 28 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 29 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 32 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 34 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 36 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 37 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 39 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 41 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 43 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 82 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 84 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| | | | | | | | | | | |
| | 101 | 111 | 121 | 131 | 141 | 151 | 161 | 171 | 181 | 191 |
| SEQ 1 | CAGTGCCCTT | CCCCGAATGA | CTGTCTCCAC | TAITCGGCAA | GATTGTAAAT | CAAGCCTGAA | GAAGCGGAGC | AATTCCTGGA | AGTCGTATGT | TCTACTGATT |
| SEQ 2 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 4 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 5 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 7 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 9 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 11 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 13 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 15 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 17 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 18 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 20 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 21 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 23 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 25 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 26 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 28 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 29 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 32 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 34 | ---AGGAAG | TTGCATGTCA | CTTGTAAGTGA | CAGGGCGTCG | TGTAAATTTT | ATAAATACCT | ATACTTGTIT | GTTCACITCT | ATGCTACTCA | TATCAATCCG |
| SEQ 36 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 37 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 39 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 41 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 43 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 82 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 84 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| | | | | | | | | | | |
| | 201 | 211 | 221 | 231 | 241 | 251 | 261 | 271 | 281 | 291 |
| SEQ 1 | TCTGTGCCTG | GCGCAGACGG | GTATATAAAT | AAAGATCACC | GCACCGAGGA | GTITCTTACC | AACCCATCAA | TAACCATCCA | CAATCTCCTA | CAACAAAAAT |
| SEQ 2 | TCTGTGCCTG | GCGCAGACGG | GTATATAAAT | AAAGATCACC | GCACCGAGGA | GTITCTTACC | AACCCATCAA | TAACCATCCA | CAATCTCCTA | CAACAAAAAT |
| SEQ 4 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 5 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 7 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 9 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 11 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 13 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 15 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 17 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 18 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 20 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 21 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 23 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 25 | TGACCCCTCT | CTTGACAACA | AAGCCGGCCA | TCCTCGCCGA | CGATTGCCTC | TACCCCGCA | TAGTCACACT | CGCACGTCCG | TTCTCCCAAC | GTCAAAACAGA |
| SEQ 26 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 28 | TGCTGTAGAT | GTGGTTGAAT | TGGTATATTA | GACCGGAGTA | CTCTATATGC | GAGAGACTAT | ACATTGAAGT | TGCCAACGTT | CTTCCAGATT | GATTAAATCAT |
| SEQ 29 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 32 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 34 | AGAAGATCAA | ACAGTCCCCT | ATACACACTT | GTCAAGACCT | ATCTATTATT | TCAAAAATCA | GCAATATGGC | TGAGACAATG | CCTAAGTGTG | AGGCAAAATGG |
| SEQ 36 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 37 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 39 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 41 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 43 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 82 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 84 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |

| | 301 | 311 | 321 | 331 | 341 | 351 | 361 | 371 | 381 | 391 |
|--------|------------|------------|-------------|------------|-------------|------------|------------|------------|-------------|------------|
| SEQ 1 | GACTGTGCCC | GATATCGAGG | TTCTCTCTGC | CGAGGGGATC | CCCTACTTCA | CTCCGGGCCA | GAACCCCTCT | GCCGGTACGG | CAGCTAACCC | CCAGACCAAT |
| SEQ 2 | GACTGTGCCC | GATATCGAGG | TTCTCTCTGC | CGAGGGGATC | CCCTACTTCA | CTCCGGGCCA | GAACCCCTCT | GCCGGTACGG | CAGCTAACCC | CCAGACCAAT |
| SEQ 4 | TGTTGTGCCT | GACATCGAGA | ACAAACCCGC | GCCGGGTATC | TCGTACTTTA | CTCCGGGCCA | AGAGCCGCCT | GCTGGCACCG | CTGCTAATCC | TCAGTCTGAT |
| SEQ 5 | TGTTGTGCCT | GACATCGAGA | ACAAACCCGC | GCCGGGTATC | TCGTACTTTA | CTCCGGGCCA | AGAGCCGCCT | GCTGGCACCG | CTGCTAATCC | TCAGTCTGAT |
| SEQ 7 | CGCCTTCGGG | TCACCAAGTC | CTCTCTCACC | CCCTACTTCA | CTCCGGGCCA | CAATGGAGGC | GCCGCCCTGC | ACCCGACGCA | CCCCAC---- | |
| SEQ 9 | GGCTCTCCCT | GACGTCGAAA | ACACCCCGCG | CGCCGGCATC | CCCTACTTTA | CACCAGCACA | GAACCCCTCT | GCTGGAACAG | CTGCCAACCC | GCAAAACAGC |
| SEQ 11 | AGTAAACCA | TCAGATGAAA | TCAAAGGTGC | TCCTGAGGTT | TCCTATTACA | CTCCAGAAC | GCCTGTTCCG | GCTGGTACTT | TTTATCCCA | ATCGTC--A |
| SEQ 13 | | | | | | | | ATGGAAA | ACAAACAATAC | TATACCG-- |
| SEQ 15 | CACCCAGAA | AAGACCTCT | CCCCCGCGCG | CCCGGGTGT | CCCTTCTTACA | CCCCGGGCCA | GGTCCCCGCC | GCCGGCACTC | CCCTCCCCCT | CACCCCC-- |
| SEQ 17 | | | | | | | | ATGGTACTT | CCACTACCTC | CGACCTC-- |
| SEQ 18 | | | | | | | | ATGGTACTT | CCACTACCTC | CGACCTC-- |
| SEQ 20 | GGCAGAAAAG | AAGACTTTGA | GCAAACCGCG | CGCCGGGGTG | CCCTACTTACA | CCCCAGCCCA | GGAGCCGCGG | GCAGGGACCC | CTTTGCAGCA | GAGGAGCG-- |
| SEQ 21 | GGCAGAAAAG | AAGACTTTGA | GCAAACCGCG | CGCCGGGGTG | CCCTACTTACA | CCCCAGCCCA | GGAGCCGCGG | GCAGGGACCC | CTTTGCAGCA | GAGGAGCG-- |
| SEQ 23 | ----ATGAC | TATTGTTAAT | GAAGGAGCCG | AAAATGTTGG | TTATTTTTACA | CTTCGCGCAA | AAATACCAGC | TGGAGCGGCG | ATAGGTGTAC | CGCAAAA-- |
| SEQ 25 | CAGCATGACG | GGCACCAGCA | ACAAGGCCGC | CCCCGGTGTG | CGGTTTATCA | CCCCGGGCCA | GGAGCCTCCC | GCGGGAACGC | CAGTCGACGC | CAGCACGG-- |
| SEQ 26 | ----ATGACG | GGCACCAGCA | ACAAGGCCGC | CCCCGGTGTG | CGGTTTATCA | CCCCGGGCCA | GGAGCCTCCC | GCGGGAACGC | CAGTCGACGC | CAGCACGG-- |
| SEQ 28 | GGCTTACGAG | ATAATCGACA | ACGTTGCGGC | TGAAGGGGTT | CCATATTACA | CACCGGCTCA | AGACCCGCCA | GCTGGTACGC | AGACAAGCGG | CTCAACG-- |
| SEQ 29 | GGCTTACGAG | ATAATCGACA | ACGTTGCGGC | TGAAGGGGTT | CCATATTACA | CACCGGCTCA | AGACCCGCCA | GCTGGTACGC | AGACAAGCGG | CTCAACG-- |
| SEQ 32 | | | | | | | | | | |
| SEQ 34 | CCATCAGAAA | ATCATCATCA | ATAAGGAAGC | TCCGAATGTT | CCCTTCTTATA | CTCCAGTGCA | AGATCCACCA | GCAGGAACGT | CTTACGATGT | TCAGCGTGAA |
| SEQ 36 | GCACGAGGG | ATTATTGACA | ACATCGCGGC | TGAAGGGGCT | CCCTACTTACA | CGCCTGCTCA | AGACYCTCCA | GCAGGCACAC | AGACCGACCG | CTCAACCA-- |
| SEQ 37 | GCACGAGGG | ATTATTGACA | ACATCGCGGC | TGAAGGGGCT | CCCTACTTACA | CGCCTGCTCA | AGACYCTCCA | GCAGGCACAC | AGACCGACCG | CTCAACCA-- |
| SEQ 39 | | | | | | | | | | |
| SEQ 41 | | | | | | | | | | |
| SEQ 43 | | | | | ATGT | CCCCACCA | CTTCGAAGCG | GCCCTGCGCG | ACCCCTCACC | GCTCGGC-- |
| SEQ 82 | AAACAAGGAG | GTGTGTCAGA | ATGTGCGTGC | CAAAGGAGTG | CAATACTTCA | ACCTGAGCA | ACTTCTGCA | CCAGGTCTCG | GTATAAACCG | TCCTCAAT-- |
| SEQ 84 | ACCGCTCTCT | GTCGACTCGA | TCGATGCAC | CAAGATCAGC | CACTTGTGTC | CACTCGAAG | TGGCCACCTT | CTCTCTGGCT | AGCTCCCGGA | CGCTCATCTG |
| | 401 | 411 | 421 | 431 | 441 | 451 | 461 | 471 | 481 | 491 |
| SEQ 1 | GG-----CC | AGAAGATCCC | CAAGCTCTTC | ACGCCCTTGA | CCATCCGTGG | CGTCACC-- | | | TTCCAGAAC | CGCCTTGGTG |
| SEQ 2 | GG-----CC | AGAAGATCCC | CAAGCTCTTC | ACGCCCTTGA | CCATCCGTGG | CGTCACC-- | | | TTCCAGAAC | CGCCTTGGTG |
| SEQ 4 | GG-----AT | CGGCACCTCC | CAAGCTCTTC | CGGCCGCTTT | CGGTGCGGGG | TCTGACC-- | | | TTTCACAAT | CGCATTGGCG |
| SEQ 5 | GG-----AT | CGGCACCTCC | CAAGCTCTTC | CGGCCGCTTT | CGGTGCGGGG | TCTGACC-- | | | TTTCACAAT | CGCATTGGCG |
| SEQ 7 | | ---GACCCC | TACGCTCTTC | CGGCCCTTAC | AAATCCGCAA | TGTGACG-- | | | CTCAAGAAC | CGCATCATG- |
| SEQ 9 | GG-----CA | ATGCCGTCCC | CAAGCTGTAC | ACACCTCTGA | CGGTGCGTGG | GGTGACC-- | | | TTCCACAAC | AGCATTTGGC |
| SEQ 11 | GA-----TG | AAGTTGCTCC | CAAAATTTTIT | CAACCTTTAA | AGATTGGTAA | GCTTGCT-- | | | TTGCCAAAC | AGAATTGGG- |
| SEQ 13 | | | GCATTTATTT | CAACCCATAA | AGATCAGTGA | CTCGATC-- | | | AC | ATTACTTAAT |
| SEQ 15 | ---G | GCGATGTCCC | TACTCTCTTC | ACCCCTCTCA | AGATCCGTGG | TGTTGAG-- | | | TTCCAGAAC | CGCTTCGCC- |
| SEQ 17 | | | AAACTCTCC | CAACCCCTCA | CCCTCCCAAA | TGGCCTT-- | | | AC | CTCCCCCAAC |
| SEQ 18 | | | AAACTCTCC | CAACCCCTCA | CCCTCCCAAA | TGGCCTT-- | | | AC | CTCCCCCAAC |
| SEQ 20 | | CCATCCC | AACGCTGTTC | AAGCCTCTGA | AGATCCGTGG | CGTCGAG-- | | | AC | CGCTCCCAAC |
| SEQ 21 | | CCATCCC | AACGCTGTTC | AAGCCTCTGA | AGATCCGTGG | CGTCGAG-- | | | AC | CGCTCCCAAC |
| SEQ 23 | | ---C | AAAATTTATTT | ACTCCTCTTA | AAATAGAGG | AGTGGAG-- | | | TTCCATAAC | AGAATTGTTT |
| SEQ 25 | | CTCC | GAGCCTCTTC | AAGCCCTTCC | GCATCCGCGA | CCTCAAC-- | | | ATCAACAAC | CGCATCTGG- |
| SEQ 26 | | CTCC | GAGCCTCTTC | AAGCCCTTCC | GCATCCGCGA | CCTCAAC-- | | | ATCAACAAC | CGCATCTGG- |
| SEQ 28 | | | AAGCTATTTC | ACACCCATCA | CCATCCGCGG | CGTCACA-- | | | TTCCCAAAC | CGCCTCTTTC |
| SEQ 29 | | | AAGCTATTTC | ACACCCATCA | CCATCCGCGG | CGTCACA-- | | | TTCCCAAAC | CGCCTCTTTC |
| SEQ 32 | | | | | | | | | | |
| SEQ 34 | GG----- | | AAGCCTATTTC | TCTCTTATTA | AAATAAGAAA | CCTGACT-- | | | CTTCAAAAC | CGGATTTTT- |
| SEQ 36 | | | AGGTTTTC | ACACBCATCA | CCATCCGAGG | CGTCACA-- | | | TTCCCAAAC | CGTCTCTTIT |
| SEQ 37 | | | AGGTTTTC | ACACBCATCA | CCATCCGAGG | CGTCACA-- | | | TTCCCAAAC | CGTCTCTTIT |
| SEQ 39 | | | | ---CCTCA | AGATCCGAGG | TCTTACC-- | | | CTCCAGAAC | CGTATTATG- |
| SEQ 41 | | | | | | | | | | |
| SEQ 43 | | ---ACGC | CGCTCAAAATA | CCCCGTCTCG | GGGCGGTTCG- | | | | GGGCCCAAC | CGGTTCTCTC |
| SEQ 82 | ---A | ATACTCTACC | AAAGGTCTTTC | ACACCCATCA | AGATTGCGCG | CATGACC-- | | | ATGCCCAAC | CGTATCTGG- |
| SEQ 84 | CCAGAGGGTG | TCAAAAAACC | GGCTTTGTTC | CAAACGTGTA | CATTGCCCTT | TGCTGCACCG | GAACAGGCGG | GTAAGATGAC | CTTCAAGAAC | CGCATCATTT |
| | 501 | 511 | 521 | 531 | 541 | 551 | 561 | 571 | 581 | 591 |
| SEQ 1 | TAAGTCCGTT | TGCCCTTGCT | CATATCGAGC | AAAGCTAATC | CCCCGTACAG- | | | | CTGCG | GGCCCTCTGC |
| SEQ 2 | | | | | | | | | CTGCG | GGCCCTCTGC |
| SEQ 4 | TGAGTGCAGT | CCAGGCAATT | ATGCTATCCA | TCCTATGCGA | GCCCTTGAT | TGGAACAGCC | GCTTACAGGG | AATGATAATG | AGTAGCTATC | GCCACTCTGC |
| SEQ 5 | | | | | | | | | CTATC | GCCACTCTGC |
| SEQ 7 | | | | | | | | | GTGTC | GCCCATGTGC |
| SEQ 9 | | | | | | | | | CTGCG | GCCCATGTGC |
| SEQ 11 | | | | | | | | | GTATC | TCCAATGTGT |
| SEQ 13 | | | | | | | | | GTGTC | ACCAATGTGC |
| SEQ 15 | | | | | | | | | GTGTC | ACCAATGTGC |
| SEQ 17 | | | | | | | | | AAAGC | CGCCATGGCC |
| SEQ 18 | | | | | | | | | AAAGC | CGCCATGGCC |
| SEQ 20 | | | | | | | | | GTCTC | GCCCATGTGC |
| SEQ 21 | | | | | | | | | GTCTC | GCCCATGTGC |
| SEQ 23 | | | | | | | | | GTGTC | GCCCATGTGC |
| SEQ 25 | | | | | | | | | GTGTC | GCCCATGTGC |
| SEQ 26 | | | | | | | | | GTGTC | GCCCATGTGC |
| SEQ 28 | | | | | | | | | GTGTC | GCCCATGTGC |
| SEQ 29 | | | | | | | | | GTGTC | GCCCATGTGC |
| SEQ 32 | | | | | | | | | GTGTC | GCCCATGTGC |
| SEQ 34 | | | | | | | | | GTGTC | GCCCATGTGC |
| SEQ 36 | | | | | | | | | GTGTC | GCCCATGTGC |
| SEQ 37 | | | | | | | | | GTGTC | GCCCATGTGC |
| SEQ 39 | | | | | | | | | GTGTC | GCCCATGTGC |
| SEQ 41 | | | | | | | | | GTGTC | GCCCATGTGC |
| SEQ 43 | | | | | | | | | GTGTC | GCCCATGTGC |
| SEQ 82 | | | | | | | | | GTGTC | GCCCATGTGC |
| SEQ 84 | | | | | | | | | GTGTC | GCCCATGTGC |

| | 601 | 611 | 621 | 631 | 641 | 651 | 661 | 671 | 681 | 691 |
|--------|------------|-------------|-------------|-------------|-------------|-------------|------------|------------|-------------|------------|
| | --2----- | ----- | ----- | ----- | ----- | ----- | ----- | 3----- | ----- | ----- |
| | ***** | ----- | ----- | ----- | ----- | ----- | ***** | ***** | ----- | ----- |
| SEQ 1 | CAATACTCCG | CC----- | ----- | ---CAGGACG | GCCACATGAC | CGAC----- | TACCACATCG | CCCATCTGGG | TGGGATCGCC | CAACCGGGAC |
| SEQ 2 | CAATACTCCG | CC----- | ----- | ---CAGGACG | GCCACATGAC | CGAC----- | TACCACATCG | CCCATCTGGG | TGGGATCGCC | CAACCGGGAC |
| SEQ 4 | CAATACTCCG | CC----- | ----- | ---GACGATG | GACACATGAC | TCCC----- | TGGCATATGG | CACATCTTGG | AGGGATTGCC | CAGCGAGGGC |
| SEQ 5 | CAATACTCCG | CC----- | ----- | ---GACGATG | GACACATGAC | TCCC----- | TGGCATATGG | CACATCTTGG | AGGGATTGCC | CAGCGAGGGC |
| SEQ 7 | ATGTACTCCT | GCGAGTCCGA | CCCGTGTCT | CCCCACGTCG | GCGCCCTAAC | AAAC----- | TACCACCTGG | CGCATCTGGG | CCACCTCGCC | CTCAAAGGCG |
| SEQ 9 | CAGTACTCCG | CA----- | ----- | ---GAAGACG | GCCACATGAC | AGAC----- | TACCACATCG | CGCACTTGGG | AGGTATTGCC | CAGCGCGGGC |
| SEQ 11 | CAATATTCTG | CT----- | ----- | GATTATAATT | TTGAAGCAAC | TCCA----- | TACCATTAA | TCCATTATGG | TTCAITAGTG | AATCGTGGG |
| SEQ 13 | ATGTATTCTG | CG----- | ---TCA | CCAACGTACA | ATCAAGCCAC | TCTG----- | TTTCATTG | TTCATTATGG | ATCAITGTCT | GTACGTGGG |
| SEQ 15 | ACCTACTCTG | CC----- | ----- | ---GACGATG | GCCACATGAC | CGAC----- | TGGCACTTGG | TCCACCTGGG | CTCTTTCGCC | CTCCCGGGTG |
| SEQ 17 | GAACAAATGG | GC----- | ----- | ---TTCGGCA | ACCACCTGCC | CAAC----- | CCCGAACTCG | CCGCGCTCTA | CGCCACCTGG | GCCTCGGGG |
| SEQ 18 | GAACAAATGG | GC----- | ----- | ---TTCGGCA | ACCACCTGCC | CAAC----- | CCCGAACTCG | CCGCGCTCTA | CGCCACCTGG | GCCTCGGGG |
| SEQ 20 | ACCTACTCAG | CC----- | ----- | ---GACGATG | GCCACCTGAC | CGAC----- | TTCCACTTGG | TGCACCTGGG | CCAGTTTCGCC | CTGCACGGCA |
| SEQ 21 | ACCTACTCAG | CC----- | ----- | ---GACGATG | GCCACCTGAC | CGAC----- | TTCCACTTGG | TGCACCTGGG | CCAGTTTCGCC | CTGCACGGCA |
| SEQ 23 | ACTTATTCCG | CT----- | ----- | GACCAAGAG | GCCACTTTGAC | AGAT----- | TTTCACCTAG | TACATCTTGG | AGCGATGGGA | ATGCGTGGG |
| SEQ 25 | CAGTACTCCG | CC----- | ----- | ---GACAAATG | GCCACGGAC | CGAC----- | TACCACCTCG | TCCACCTGGG | CCAGTTTCGCC | CTGCACGGG |
| SEQ 26 | CAGTACTCCG | CC----- | ----- | ---GACAAATG | GCCACGGAC | CGAC----- | TACCACCTCG | TCCACCTGGG | CCAGTTTCGCC | CTGCACGGG |
| SEQ 28 | CAATACTCCG | CC----- | ----- | ---AAAGATG | GTTATGCCAC | TGAT----- | TGGCACTTGA | CTCACCTCGG | GGGAATAATC | CAAGAGGGC |
| SEQ 29 | CAATACTCCG | CC----- | ----- | ---AAAGATG | GTTATGCCAC | TGAT----- | TGGCACTTGA | CTCACCTCGG | GGGAATAATC | CAAGAGGGC |
| SEQ 32 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 34 | CAATATTCTG | CA----- | ----- | ---AAGGATG | GTGTCATGAC | CCCC----- | TGGCACAAC | AACACCTGGG | CAGCTTCGCA | GCACGAGGTC |
| SEQ 36 | CAATACTCCG | CC----- | ----- | ---AAAGATG | GATATGCTAC | TGAT----- | TGGCACTTGA | CTCATCTCGG | AGGCATTATC | CAACGAGGCG |
| SEQ 37 | CAATACTCCG | CC----- | ----- | ---AAAGATG | GATATGCTAC | TGAT----- | TGGCACTTGA | CTCATCTCGG | AGGCATTATC | CAACGAGGCG |
| SEQ 39 | CAGTACTCTG | CT----- | ----- | ---CCGACG | GACACTACAC | AATG----- | TGGCATCACA | CCACATATGG | CGGCATATC | CAACCGGGTC |
| SEQ 41 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 43 | GAGGGCCTGG | CG---ACGTT | TGACGAGGCG | GACCCGTCCA | AGCGCGGCAT | CCCGACGGAG | CAGCTGGTGG | AGCTGTACCG | CGCTGGGGG | CAGGGCGAGT |
| SEQ 82 | CAATACAGTG | CC----- | ----- | ---CGTAGC | GCTTTACAGC | GCCT----- | TGGCACTTGG | CCCACTACGG | CGCACTGGCC | CAACGTGGCC |
| SEQ 84 | CAGTACTCTG | CG----- | ----- | ---AACAAATG | GTCTTCTCTAC | TCCG----- | TACCACATTG | CGCATTTGGG | ATCGTTTGGC | CTGCACGGTG |
| | 701 | 711 | 721 | 731 | 741 | 751 | 761 | 771 | 781 | 791 |
| | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 1 | CGGCGCTGAT | GCTGATTGAG | GCGACCGCCG | TCCAGCCCGA | A---GGCCGC | ATCACCCCTC | AGGATGTGCG | TCTGTGGAAG | GACTTCC--- | -----CA |
| SEQ 2 | CGGCGCTGAT | GCTGATTGAG | GCGACCGCCG | TCCAGCCCGA | A---GGCCGC | ATCACCCCTC | AGGATGTGCG | TCTGTGGAAG | GACTTCC--- | -----CA |
| SEQ 4 | CAGGATTTCT | GATGGTCGAG | GCAACAGCAG | TGCAACCCGA | A---GGCAGG | ATCACCCCGC | AGGACCTGGG | ACTATGGAAG | GACTCG--- | -----CA |
| SEQ 5 | CAGGATTTCT | GATGGTCGAG | GCAACAGCAG | TGCAACCCGA | A---GGCAGG | ATCACCCCGC | AGGACCTGGG | ACTATGGAAG | GACTCG--- | -----CA |
| SEQ 7 | CAGGCTCTGT | CTTCATCGAA | GCCACCGCCG | TGCAACCCGA | A---GGCAGG | ATCTCCCGCA | ACGACTCGGG | CCTCTGGCAG | GACGCGACCA | CCTCGGAACA |
| SEQ 9 | CGGCTCTGT | GATGATCGAG | GCAACCTCCG | TCTCACCTGA | A---GGCAGG | ATCACCCCGC | AGGACCTGGG | TTTATGGAAG | GACTCG--- | -----CA |
| SEQ 11 | CAGGTATCAC | CAITTTTAGAG | AGCAACGGCTG | TTTCTCCTGA | G---GGTGGA | TTATCACCTC | ATGATTATAG | AACTGGAAG | GATGAA--- | -----CA |
| SEQ 13 | CAGCATTAAT | CAITTTTAGAG | AGTATCTTTG | TGTCGGAAGA | T---TCCGGA | TTATCACCTC | ATGATTATAG | TTTGTGGAAT | GATGAT--- | -----CA |
| SEQ 15 | TCCCGCTCAC | CATCTTCGAG | GCCACCGCCG | TCTTCCCGAA | C---GGCCGC | ATCACCCCGC | AGTGTCTGCG | TCTCTGGCAG | GACTTC--- | -----CA |
| SEQ 17 | ACTGGGGCCT | GATTCTCACC | GGCAACGTCC | AAGTCGACCA | CGCGACAAG | GCGGACGCCC | ACGACATCAG | CCCCAACCA | CCCGGACCA | CGCCGAGCA |
| SEQ 18 | ACTGGGGCCT | GATTCTCACC | GGCAACGTCC | AAGTCGACCA | CGCGACAAG | GCGGACGCCC | ACGACATCAG | CCCCAACCA | CCCGGACCA | CGCCGAGCA |
| SEQ 20 | CGGCGCTGAC | CATGTCGAG | GCCACATCCG | TCAACGCCAA | C---GGACGC | ATCTCGCCCG | AGGACAGCGG | CCTGTGGCAA | GACAGC--- | -----CA |
| SEQ 21 | CGGCGCTGAC | CATGTCGAG | GCCACATCCG | TCAACGCCAA | C---GGACGC | ATCTCGCCCG | AGGACAGCGG | CCTGTGGCAA | GACAGC--- | -----CA |
| SEQ 23 | CTGCGCTTGT | AATGGTAGAA | GCGACAGCGG | TTTCCCGAGA | G---GGACGC | ATTTACCTTA | ATGATTGAGT | ATTATGGAAT | GAGTCG--- | -----CA |
| SEQ 25 | CGGCGCTGTC | CATGGTCGAG | GCCACCGCCG | TGAGGCTCTG | T---GGCCGC | ATCTCCCGCG | AGGATGTGCG | TTTGTGGCAG | GACTCG--- | -----CA |
| SEQ 26 | CGGCGCTGTC | CATGGTCGAG | GCCACCGCCG | TGAGGCTCTG | T---GGCCGC | ATCTCCCGCG | AGGATGTGCG | TTTGTGGCAG | GACTCG--- | -----CA |
| SEQ 28 | CGGATTTGTC | CATGGTGGAG | GCTACCGCTG | TACAAAACCA | C---GGTCGC | ATCACACCTC | AGGATGTTGG | TCTGTGGGAA | GACGGC--- | -----CA |
| SEQ 29 | CGGATTTGTC | CATGGTGGAG | GCTACCGCTG | TACAAAACCA | C---GGTCGC | ATCACACCTC | AGGATGTTGG | TCTGTGGGAA | GACGGC--- | -----CA |
| SEQ 32 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 34 | CGGCTCTCAT | TGTCACAGAA | GTCAACGCAG | TTTCAACAGA | G---GGACGC | ATCAGTCTCT | AGGATGCAAG | CATCTACGAT | GATGGG--- | -----CA |
| SEQ 36 | CGGACTGTCT | CATGGTAGAG | GCCACCGCTG | TTCAAAACCA | C---GGTCGC | ATCACGCTCT | AGGACGTTGG | TCTCTGGGAA | GATGGA--- | -----CA |
| SEQ 37 | CGGACTGTCT | CATGGTAGAG | GCCACCGCTG | TTCAAAACCA | C---GGTCGC | ATCACGCTCT | AGGACGTTGG | TCTCTGGGAA | GATGGA--- | -----CA |
| SEQ 39 | CGGACTGTCT | CTGGCTTGAA | GCCACAGCGC | TGACTCTCTA | A---GGTCGC | ATCACGCTCT | AAGACGTGCG | TATCTGGCAA | GATTCT--- | -----CA |
| SEQ 41 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 43 | GGGCGCAGAT | CCAGACGGG | AACGTCTAGA | TGCAACCGGA | GCACCTCGAG | CGCCCGGGCA | ACATGTTGGT | GCCGCGCGAC | GCCGAGCCT | CGGGCGAGCG |
| SEQ 82 | CTGCGCTCAT | CATGCTAGAA | GCTACCGCAG | TTCAAGCAGC | T---GGCCGT | ATCACACCTG | AAGATTCTGG | CATCTGGCTA | GACTCT--- | -----CA |
| SEQ 84 | TGGGAAACGT | CATGGTCGAA | GCATCTGGTG | TTGAGCCAGA | G---GGGAGG | ATCACGCTCT | AGGACCTGGG | TATTTGGTGC | AACAGC--- | -----CA |
| | 801 | 811 | 821 | 831 | 841 | 851 | 861 | 871 | 881 | 891 |
| | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 1 | GATCGCCCGG | ---ATGCGCC | GGGTCTACGA | CTTCGTGAC | AGCCAGGGC | CAGAAGATCG | GCGTG----- | ---CAGCTT | GCCCATGCGG | GCCGGAAGC |
| SEQ 2 | GATCGCCCGG | ---ATGCGCC | GGGTCTACGA | CTTCGTGAC | AGCCAGGGC | CAGAAGATCG | GCGTG----- | ---CAGCTT | GCCCATGCGG | GCCGGAAGC |
| SEQ 4 | GATTGAGCCA | ---TTGAGCC | CGGTGATCGA | GTTTGTCCAC | AGTCAGAAC | CAGCTTATCG | GCGTG----- | ---CAGATC | GCACACGCG | GTCCGAAGG |
| SEQ 5 | GATTGAGCCA | ---TTGAGCC | CGGTGATCGA | GTTTGTCCAC | AGTCAGAAC | CAGCTTATCG | GCGTG----- | ---CAGATC | GCACACGCG | GTCCGAAGG |
| SEQ 7 | ATTCTCTGGG | ---CTGAAGC | GGGTCTGCGA | GTTTCATGAC | GCACAGGGC | CAGCAAGGTC | GGATC----- | ---CAGCTT | GCAGATGCG | GCCGGAAGC |
| SEQ 9 | GATTGCGCCC | ---ATGAAGC | CGGTCTGCGA | CTTCGTGAC | TGCGACTCC | CAGAAGATTG | GCGTG----- | ---CAGATT | GCCACGCGG | GCCGGAAGG |
| SEQ 11 | AGCAGAGAAA | ---TTGAAGC | CAATTGTGCGA | TTACGCTCAT | TCTCAAAAG | CAATTAAATTG | CCATC----- | ---CAATTG | GGCCATGGTG | GTAGAAAAGC |
| SEQ 13 | AGCTCACAGT | ---TTACGGA | AAATTGTGTA | TTTATTATCAT | GATCAAGAC | GGAAATTGTCT | GTATA----- | ---CAATTG | AATCACGCTG | GGCGAAGAT |
| SEQ 15 | GATTGCGCCC | ---CTCAAGC | GCATCGTCTGA | CTACATGAC | TCCACAGGGC | CAGAAGGCGG | GTATC----- | ---CAGCTT | GCCACGCGG | GCCGGAAGG |
| SEQ 17 | GACCGTCACG | GCCTTCAAGG | CCTGGGCGGA | CGCCGCGCGC | CTGAATGGC | CAGTCCAAAA | CGCTGTGGT | CGTGCAGATC | AACCAACCTG | GTCCGCAAG |
| SEQ 18 | GACCGTCACG | GCCTTCAAGG | CCTGGGCGGA | CGCCGCGCGC | CTGAATGGC | CAGTCCAAAA | CGCTGTGGT | CGTGCAGATC | AACCAACCTG | GTCCGCAAG |
| SEQ 20 | GATCGCTCTC | ---CTGCGCC | GCATCGTCTGA | CTACGTGAC | AGCCAGGGC | CAAAAGATCG | CCATC----- | ---CAACTG | GCTCATGCGG | GCCGCAAGG |
| SEQ 21 | GATCGCTCTC | ---CTGCGCC | GCATCGTCTGA | CTACGTGAC | AGCCAGGGC | CAAAAGATCG | CCATC----- | ---CAACTG | GCTCATGCGG | GCCGCAAGG |
| SEQ 23 | AATGAAGCGG | ---TTACGAA | GAATTGTGTA | ATTGTCTCAT | TGCGAAAAAT | CAAAAAATTG | GGATT----- | ---CAATTG | GGCCATGCTG | GTAGAAAAGC |
| SEQ 25 | GATTGCGCGG | ---CTGAAGC | GCATCGTCTGA | CTTTATCCAT | TGCGAGAAC | CAGGTTCGCG | CCATC----- | ---CAGCTC | GCCACGCGG | GTCCGAAGG |
| SEQ 26 | GATTGCGCGG | ---CTGAAGC | GCATCGTCTGA | CTTTATCCAT | TGCGAGAAC | CAGGTTCGCG | CCATC----- | ---CAGCTC | GCCACGCGG | GTCCGAAGG |
| SEQ 28 | GATCGAGCCT | ---CTGAAGC | GCATCACCAC | TTTCGCGCAC | AGTCAGAGC | CAGAAAATTG | GTATC----- | ---CAGCTG | TGCGATGCGG | GTCCGAAGG |
| SEQ 29 | GATCGAGCCT | ---CTGAAGC | GCATCACCAC | TTTCGCGCAC | AGTCAGAGC | CAGAAAATTG | GTATC----- | ---CAGCTG | TGCGATGCGG | GTCCGAAGG |
| SEQ 32 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 34 | GCTTGGACCT | ---CTCGGG | ATATTGTGGA | CTTTGTACAC | AGCCAGGGC | GCCAAGATTG | CTATT----- | ---CAGATA | GGTCATGCTG | GGAGAAAAGC |
| SEQ 36 | AATCGAGCCC | T---TTGAAGC | GCATCACTAC | TTTGTGCCAC | AGCCAAAAGC | CAGAAGATTG | GTAT----- | ---TCAGCTC | TGCGACGCTG | GTCTTAAGGC |
| SEQ 37 | AATCGAGCCC | ---TTGAAGC | GCATCACTAC | TTTGTGCCAC | AGCCAAAAGC | CAGAAGATTG | GTAT----- | ---TCAGCTC | TGCGACGCTG | GTCTTAAGGC |
| SEQ 39 | GATCGAGCCT | C---TTGCCAA | GGTCGTC-GA | GTTCGCCAC | TCCAGAAC | CAGAAGATCA | TGATT----- | ---CAGTTG | GGCATGCGG | GCCGCAAGG |
| SEQ 41 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 43 | CTTCGACATG | TTTTCGAAGC | TGCGCGCGCG | CGCCAAAGAG | CACGGCAGC | CTC-ATCGTC | GCG----- | ---CAGGTC | GGACACCCCG | GTCCGCAAGG |
| SEQ 82 | TGTTGAGGGA | ---CTGCGAA | AGCACGTCTGA | TTTGTGCCAC | GCCTTACAG | TCTCTTATCG | GTATC----- | ---CAGATT | GGCCATGCTG | GTCCGAAGG |
| SEQ 84 | TGCGGATGCA | ---CACAAAG | CGCTGGTGTG | GGTGCTCAAG | TGCTTACAG | GATGGTCTCG | GTGTA----- | ---GGGCTG | CAACTGCGCG | ATGCGGGAAG |

| | 1201 | 1211 | 1221 | 1231 | 1241 | 1251 | 1261 | 1271 | 1281 | 1291 |
|--------|------------|------------|------------|------------|------------|------------|------------|-------------|------------|------------|
| SEQ 1 | GGATACCTGC | TGTGTCATT | CCTCTCGCG | GCGGCCAAC | | | | | | |
| SEQ 2 | GGATACCTGC | TGTGTCATT | CCTCTCGCG | GCGGCCAAC | | | | | | |
| SEQ 4 | GGCTATCTTC | TGATGTCGTT | CCTCTCCCT | GCGGTCAAC | | | | | | |
| SEQ 5 | GGCTATCTTC | TGATGTCGTT | CCTCTCCCT | GCGGTCAAC | | | | | | |
| SEQ 7 | GGCTATCTCA | TCAACGAGTT | CCTGAGCCG | GTACGGAAT | | | | | | |
| SEQ 9 | GGGTATCTTC | TCTGTCCTT | CCTATCACC | TCTTCCAAC | | | | | | |
| SEQ 11 | GGTTATTGTA | TAAATGAGTT | CTATAGTCT | ATTTCAAAC | | | | | | |
| SEQ 13 | GGATGTTTAA | TACACCAATT | TTTAAAGTAA | TTGACAAAC | | | | | | |
| SEQ 15 | GGCTACCTCA | TTTCCGAGTT | CTTGAGCCCC | ATCTCCAAC | | | | | | |
| SEQ 17 | GGATACCTGT | TGGCGCAGTT | CTTGAGCAAG | AAGACAAAC | | | | | | |
| SEQ 18 | GGATACCTGT | TGGCGCAGTT | CTTGAGCAAG | AAGACAAAC | | | | | | |
| SEQ 20 | GGTTACCTGA | TCACCGAGTT | CCTTTCGCC | CTATCAAAC | TAAGTGGAGA | TACTTGTGT | GGGGTGTGC | GCATACTCCC | TCGGGTGTGA | CTTCTATTAA |
| SEQ 21 | GGTTACCTGA | TCACCGAGTT | CCTTTCGCC | CTATCAAAC | | | | | | |
| SEQ 23 | GGTTATCTTA | TATCGTCAAC | AGTTAGTCT | GCCACTAAT | | | | | | |
| SEQ 25 | | | | | | | | | | |
| SEQ 26 | | | | | | | | | | |
| SEQ 28 | GGATATCTAC | TGCATCAATT | CTTGAGTCCG | GTAAGCAAT | | | | | | |
| SEQ 29 | GGATATCTAC | TGCATCAATT | CTTGAGTCCG | GTAAGCAAT | | | | | | |
| SEQ 32 | | | | | | | | | | |
| SEQ 34 | | | | | | | | | | |
| SEQ 36 | GGATACKTGC | TTACCCAGTT | CTTGAGTCCA | GTCAGTAAC | | | | | | |
| SEQ 37 | GGATACKTGC | TTACCCAGTT | CTTGAGTCCA | GTCAGTAAC | | | | | | |
| SEQ 39 | GGATACCTCC | TCCACGAAAT | CATCTGCCTG | AGAGCAACA | | | | | | |
| SEQ 41 | GGGTACCTCA | TCCACGAAAT | CCTCTCACC | ATTACCAAC | | | | | | |
| SEQ 43 | GGTTACCTGC | TGGCCCAATT | CCTGTCGGA | ACAACCAAC | | | | | | |
| SEQ 82 | GGTTATCTTG | TTTCCAGCTT | CCTGTCGCT | GCCACCAAC | | | | | | |
| SEQ 84 | GGATACCTGA | TGCACTCGTT | CCTCAGCCG | TTGACCAAT | | | | | | |
| | | | | | | | | | | |
| | 1301 | 1311 | 1321 | 1331 | 1341 | 1351 | 1361 | 1371 | 1381 | 1391 |
| SEQ 1 | | | AACCGCAC | GGACCACTAG | GGCGGTCGT | TCGAGAACC | CATCCGCTG | TCTCTCGAGA | TTGCGCAGTT | GAATCGGGAC |
| SEQ 2 | | | AACCGCAC | GGACCACTAG | GGCGGTCGT | TCGAGAACC | CATCCGCTG | TCTCTCGAGA | TTGCGCAGTT | GAATCGGGAC |
| SEQ 4 | | | ACGAGAAC | AGACGAGTAC | GGAGGCACT | TTGAGAATCG | CATCCGCTC | AGTCTCGAGA | TCGCCAAGCT | CACCCGCGAA |
| SEQ 5 | | | ACGAGAAC | AGACGAGTAC | GGAGGCACT | TTGAGAATCG | CATCCGCTC | AGTCTCGAGA | TCGCCAAGCT | CACCCGCGAA |
| SEQ 7 | | | AAGCGGAC | GGATGCGTAC | GGCGGAGCT | TTGAGAACC | GACCCGCTC | GTGCGCGAGG | TTGCGCGGCG | TATTCGTGCG |
| SEQ 9 | | | ACGCGCAC | CGACGAGTAC | GGCGGTCGT | TTGAGAACC | CATCCGCTC | TCTCTCGAAA | TCGCCAAGCT | CACCCGCGAA |
| SEQ 11 | | | AAGAGAAC | AGATGAATAC | GGTGGCAGTT | TTGAAAATAG | AACCCAGATT | TTAAAGGAAG | TTATCGATAG | TGTTAAATCA |
| SEQ 13 | | | AAGAGAGC | TGACCAATAC | GGGGGCTCAT | TTGAAAACAG | AGTTAGATT | CTTTTACAAA | TAATTGAGAA | TATAAAACGA |
| SEQ 15 | | | CAGCGTAC | CGACGAGTAC | GGTGGCTCT | TCGAGAACC | CACCCGCTG | CTCCGCGAGA | TCATCTCGGC | CGTCCGCTCC |
| SEQ 17 | | | AGGCGCGG | GGATGAGTAT | GGCGGTCGT | TCGAGAACC | GCGGAGGATT | GTGCGGAGG | TTATTAAGGA | GTGCGAGGAG |
| SEQ 18 | | | AGGCGCGG | GGATGAGTAT | GGCGGTCGT | TCGAGAACC | GCGGAGGATT | GTGCGGAGG | TTATTAAGGA | GTGCGAGGAG |
| SEQ 20 | CATTTTATTT | CCTGGCACCG | AGAAACGGAC | AGACAAGTAC | GGCGGAGCT | TTGAGAACC | CACCCGCTC | CTGATCGATA | TTATCAAGGC | CGTCCGGGCA |
| SEQ 21 | | | AAACGGAC | AGACAAGTAC | GGCGGAGCT | TTGAGAACC | CACCCGCTC | CTGATCGATA | TTATCAAGGC | CGTCCGGGCA |
| SEQ 23 | | | GACCGCAA | TGACAAGTAT | GGTGGGACAT | TTGAGAAACG | TATTTTGT | CCTATGGAAG | TTGTCCATT | TGTTCTGATA |
| SEQ 25 | | | | | | | | | | |
| SEQ 26 | | | | | | | | | | |
| SEQ 28 | | | CAAAGAAC | CGACGAGTAT | GG | | | | | |
| SEQ 29 | | | CAAAGAAC | CGACGAGTAT | GG | | | | | |
| SEQ 32 | | | AAC | CGACGAGTAT | GGTGGCAGTT | TCGAGAACC | TATCAGAGTT | GTCTTGGAAG | TGCTTGACCT | CATCCGCGCT |
| SEQ 34 | | | | | | | | | | |
| SEQ 36 | | | CAAAGAAC | CGATGAGTAT | GGTGGCAGCT | TCGAGAACC | TATCAGAGTT | GTCTTGGAAG | TCATTG | |
| SEQ 37 | | | CAAAGAAC | CGATGAGTAT | GGTGGCAGCT | TCGAGAACC | TATCAGAGTT | GTCTTGGAAG | TCATTG | |
| SEQ 39 | | | CCAGGACC | GACAAGTACG | GGGGAAGCT | GGGAAAACCG | CATCTGCTG | ACAATGGAAA | TCGCTCGACC | TTGTCCGCG |
| SEQ 41 | | | CGCGGAC | AGATTCTTAC | GGCGGTCCT | TCGAAAACCG | TACCCGCTA | CTCATTTGAAA | TCGTAACAGC | CGTCCGAGCC |
| SEQ 43 | | | CAGCGCAC | CGACGAGTAC | GGCGGAGCT | TCGAAAACCG | CATCCGCTG | ATCCTCGAGG | TCACGGCCGA | GGTCCGCGAG |
| SEQ 82 | | | AAGCGTAC | CGACAAGTAC | GGAGGTAGCT | TCGAGAACC | AGTGGCCTT | GCTCTCGAGA | TTGTGAGGC | TGCACGAGCT |
| SEQ 84 | | | CAGCGTAC | CGACGAGTAC | GGCGGAGCT | TCGAGAACC | CGCTCGATT | CTGCTCAACG | TTGCCGCTG | AATCCGCGCA |
| | | | | | | | | | | |
| | 1401 | 1411 | 1421 | 1431 | 1441 | 1451 | 1461 | 1471 | 1481 | 1491 |
| SEQ 1 | GCGGTCGGCC | CTCATGTGCC | | | GTCTTT | CCTGCGCAAT | TCGGGCTCGG | ACTGGTGCGA | GGAGACCTG | CCGGA |
| SEQ 2 | GCGGTCGGCC | CTCATGTGCC | | | GTCTTT | CCTGCGCAAT | TCGGGCTCGG | ACTGGTGCGA | GGAGACCTG | CCGGA |
| SEQ 4 | AATGTGCCCA | AGGATATGCC | | | GTCTTT | CCTGCGGCT | TCCGCCACCG | ATTGGCTGGA | GGAGGTGCAG | CCGAA |
| SEQ 5 | AATGTGCCCA | AGGATATGCC | | | GTCTTT | CCTGCGGCT | TCCGCCACCG | ATTGGCTGGA | GGAGGTGCAG | CCGAA |
| SEQ 7 | GTGATTCCCG | AGGGGATGCC | | | CTGTT | TCTGCGTATC | AGCGCCACGG | AGTGGTTGGA | GGGTGAGCCG | GTGGC |
| SEQ 9 | GCGGTCGGCC | CCAACGTGCC | | | GTCTTT | TCTGCGTATC | TCCGCCACCG | ACTGGATCGA | GGAGACCTCT | CCGGA |
| SEQ 11 | AGTATTCCAA | ACGATGTGCC | | | GTGTT | TTTGAGAATC | TCTGCTGCTG | AAAATAGTCC | TGATCCA | |
| SEQ 13 | AAGATAGAAA | CA | CC | | ATTTT | CTTAAAGTTT | CCAATGTGAG | ATAATTGTAG | TGATCCG | |
| SEQ 15 | GTATCTCCCG | AGGACATGCC | | | CTCTT | CGTCCGTGTC | TCCGCCACCG | AGTGGATGGA | GTACACC | |
| SEQ 17 | CAGGTGACTG | AGGCGGTGGG | TGAAGAGGAG | GCGAAGAAAT | TTGTGTTGGG | AATCAAGCTG | AACAGTGCAG | ATTGGCAGGC | GGGACGCGAT | GGA |
| SEQ 18 | CAGGTGACTG | AGGCGGTGGG | TGAAGAGGAG | GCGAAGAAAT | TTGTGTTGGG | AATCAAGCTG | AACAGTGCAG | ATTGGCAGGC | GGGACGCGAT | GGAAAG |
| SEQ 20 | GTGATTCCCG | AGGAGATGCC | | | CTCTT | CGTCCGAATC | TCCGCCACCG | AATGGATGGA | GTACGCCGCG | |
| SEQ 21 | GTGATTCCCG | AGGAGATGCC | | | CTCTT | CGTCCGAATC | TCCGCCACCG | AATGGATGGA | GTACGCCGCG | |
| SEQ 23 | GCAATTCCAG | ATAGTATGCC | | | TTGTT | TTATAGAGTA | ACGGCTACAG | ATTGGTTGCC | CAAAGGACAA | |
| SEQ 25 | | | | | | | | | | |
| SEQ 26 | | | | | | | | | | |
| SEQ 28 | | | | | | | | | | |
| SEQ 29 | | | | | | | | | | |
| SEQ 32 | GCCATCCCCG | AAACTACACC | T | | GTCTT | CGTTCGTGTC | AGTGCAACTG | ATTGGTTGCA | GTTTGACTCT | CAATTCAAAG |
| SEQ 34 | | | | | | | | | | |
| SEQ 36 | | | | | | | | | | |
| SEQ 37 | | | | | | | | | | |
| SEQ 39 | CATT | | | | | | | | | |
| SEQ 41 | GCGATGCCCT | CCAGCATGCC | T | | CTCTT | CCTCCGCTC | TCTCTACAG | AATGGATGGA | AGATACCGAC | ATCGGC |
| SEQ 43 | CGGACGAGCA | AGAATTTCAT | C | | CTCGG | CATCAAAAT | AACAGCGCTG | AGTTCCAGGA | GAAG | |
| SEQ 82 | GTTATGCTTG | AGGACATGCC | | | TTGTT | CATCTGCATC | AGTGGAACTG | ACTGGCTGGA | GAACAAACCT | GAG |
| SEQ 84 | GAATTCCCCA | ACAAGGGT | | | CTCTG | GGTGGCGCTG | AGCTCCACCG | ACTGGGCCGA | CCAAGCGCAC | CAA |

| | 1501 | 1511 | 1521 | 1531 | 1541 | 1551 | 1561 | 1571 | 1581 | 1591 |
|--------|------------|--------------|------------|------------|-------------|------------|-------------|------------|------------|-------------|
| SEQ 1 | ----- | GCAGAGCTGG | AAGTCGGAGG | ATACCGTGCG | GTTTCGCGCAG | GAGCTGGTCA | AGCAGGGGCGC | CGTTGATCTG | ATCGATATCA | GCAGCGGTGG |
| SEQ 2 | ----- | GCAGAGCTGG | AAGTCGGAGG | ATACCGTGCG | GTTTCGCGCAG | GAGCTGGTCA | AGCAGGGGCGC | CGTTGATCTG | ATCGATATCA | GCAGCGGTGG |
| SEQ 4 | ----- | CAAGCCAGCTGG | CGAGGCGTGG | ACACTGTCCG | ATTTCGCGAAG | ATCCTGGCAG | AAACGGGTGA | CGTTGACGTG | CTTGACGTGA | GCAGTGGCGG |
| SEQ 5 | ----- | CAAGCCAGCTGG | CGAGGCGTGG | ACACTGTCCG | ATTTCGCGAAG | ATCCTGGCAG | AAACGGGTGA | CGTTGACGTG | CTTGACGTGA | GCAGTGGCGG |
| SEQ 7 | ----- | CGCGGAGTGC | GCGGAGCTGG | GATAT---GC | AGAGCTCGCT | GGAGCTGCTC | CCGAATGGGG | CATTGACCTG | GTGGATGTCA | GCTCCGCGCG |
| SEQ 9 | ----- | GGAATCTGTG | AAGCTCTCTG | ACTCCGTCCG | CTTCGCCGAA | GCCCTCGCTG | CCCAGGGGCGC | TATTGACCTG | ATCGACGTCT | CTTCGCGCGG |
| SEQ 11 | ----- | GAAGCTTGG | ACTATTGAAG | ATTCCAAAA | --AATTAGCT | GACATTTTAG | TAGAAAAGGG | TATTGCTTTG | GTTGATGTTT | CATCTCGGTGG |
| SEQ 13 | ----- | GAAGCTTGG | TCTACGGAAG | ATGCAATTGA | --AGTTGGCC | GATCTTGTTA | TTGATTTAGG | AGTAAAGGTG | ATCGACGTGA | CATCAGGTGG |
| SEQ 15 | ----- | GGCCA | GCCCTCGTGG | GACCTCCAGC | AGACCAATTG | --AGCTCGCC | CCGACCTCGG | CGTCGACCTC | CTCGACGTCT | CTTCGCGCGG |
| SEQ 17 | AGGAGGAGGA | GGAGACGGAT | ACGGCGGAGG | AGGTGTTGA | --AGCAGATT | GAGCTTTTGG | AGCAGTGGGG | GATCGACTTT | GTGAGGTGTA | GCAGTGGCGG |
| SEQ 18 | --GAGGAGGA | GGAGACGGAT | ACGGCGGAGG | AGGTGTTGA | --AGCAGATT | GAGCTTTTGG | AGCAGTGGGG | GATCGACTTT | GTGAGGTGTA | GCAGTGGCGG |
| SEQ 20 | ----- | GA | GCCTAGCTGG | GACCTCGAGC | AGAGCACAC | --AGCTTGCC | AAGCTCTCTC | CGGACCTGGG | TGTCGACCTG | CTCGACGTCA |
| SEQ 21 | ----- | GA | GCCTAGCTGG | GACCTCGAGC | AGAGCACAC | --AGCTTGCC | AAGCTCTCTC | CGGACCTGGG | TGTCGACCTG | CTCGACGTCA |
| SEQ 23 | ----- | GGATGG | GAGATAGAAG | ATACAGTTG | --CATTAGCA | GCGAGGCTTC | GCGATGGTGG | TGTTGACTTG | ATAGATGTGA | GCTCTGGTGG |
| SEQ 25 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 26 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 28 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 29 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 32 | ACGAGTTTCC | TGAAAGCTGG | ACAGTCGAGC | AGACTT---G | TCAACTCGCG | CGTATCTTGC | CCAAGCATGG | AGTAGACTTG | GTGGAGCTCA | GCTCAGGCGG |
| SEQ 34 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 36 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 37 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 39 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 41 | --AAGAAGTT | CGGAAGCTGG | GATGTCGAAA | GCACGATCA | --AGATCTCC | AAATCTCTGG | CCGACTTGGG | CGTTGATCTC | CTCGACGTGT | CTTCGCGGTGG |
| SEQ 43 | ----- | GGTTTCAAG | CCA---GAGG | AGGCGGTGC | --AGTTGTGC | GAGGCCCTCG | AGGCCGCGGG | CATGGATTTT | GTGAGACGGA | GCAGCGGTGG |
| SEQ 82 | --TACGAGGG | AGAGACCTGG | ACTCTTGAGC | AGAGCATCA | --AGCTTGCA | CACCACTTAG | CAGACCGTGG | TGTCGATGTT | TTGGATGTTT | CCAATGGTGG |
| SEQ 84 | ----- | GC | CGACTCTTGG | ACCGTTGACC | AGACGGTTG | --AACTCGCC | AAGATGCTCC | AAGAGGCTCG | AGTCGACCTG | GCTCCGCGCG |
| SEQ 1 | 1601 | 1611 | 1621 | 1631 | 1641 | 1651 | 1661 | 1671 | 1681 | 1691 |
| SEQ 1 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 2 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 4 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 5 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 7 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 9 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 11 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 13 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 15 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 17 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 18 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 20 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 21 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 23 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 25 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 26 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 28 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 29 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 32 | TATCCATCCT | AAG----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 34 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 36 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 37 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 39 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 41 | GAATCATCCT | CAG----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 43 | CTATGAGAGT | TTT----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 82 | CATCCACAAG | ATG----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 84 | CCTGGTTCCA | TTC----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 1 | 1701 | 1711 | 1721 | 1731 | 1741 | 1751 | 1761 | 1771 | 1781 | 1791 |
| SEQ 1 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 2 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 4 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 5 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 7 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 9 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 11 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 13 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 15 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 17 | ATACAGATGG | CCAACCGTCC | CAAGCCCGAA | AGTCCGAAAC | GCACCATGGC | CCGCGAGGCC | TTCTTCTCTG | AGTTCGCCAA | GATCATCCGC | ACCAAG---T |
| SEQ 18 | ----- | ATGG | CCAACCGTCC | CAAGCCCGAA | AGTCCGAAAC | GCACCATGGC | TTCTTCTCTG | AGTTCGCCAA | GATCATCCGC | ACCAAG---T |
| SEQ 20 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 21 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 23 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 25 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 26 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 28 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 29 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 32 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 34 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 36 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 37 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 39 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 41 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 43 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 82 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 84 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |

| | 1801 | 1811 | 1821 | 1831 | 1841 | 1851 | 1861 | 1871 | 1881 | 1891 | |
|--------|------------|------------|-------------|------------|------------|---------|------------|------------|------------|-------------|-------------|
| SEQ 1 | --- | AAGCT | GCTGGTTGCC | GCCGTGGGTG | CCATCACC | --- | --- | --- | --- | AACG | GCAAGCAGGC |
| SEQ 2 | --- | AAGCT | GCTGGTTGCC | GCCGTGGGTG | CCATCACC | --- | --- | --- | --- | AACG | GCAAGCAGGC |
| SEQ 4 | --- | AAACT | CGCAGTGGCA | TCAGTGGGTA | TGATTGCC | --- | --- | --- | --- | ACGC | CGCATTTTGCC |
| SEQ 5 | --- | AAACT | CGCAGTGGCA | TCAGTGGGTA | TGATTGCC | --- | --- | --- | --- | ACGC | CGCATTTTGCC |
| SEQ 7 | --- | GCTGG | CGCGTCGACT | CTTGTGGGTG | CTGTAGGTCT | GATCACC | CGAT | TCGGAACAGG | CGAGGGGACT | AGTTTCAGGGA | GCGGCAGCAGG |
| SEQ 9 | --- | AAGCT | CCTTGTGTGCG | ACGGTGGGCA | CGATCACG | --- | --- | --- | --- | AACG | GTAAGCAGGC |
| SEQ 11 | --- | AAGTT | AATTGGTCAGT | TGCGTTGGTG | GGCTTGAA | --- | --- | --- | --- | A | AAGATCCTGA |
| SEQ 13 | --- | CGATG | TTTGATCTGCA | TGCAGTGGAG | GATTAGAT | --- | --- | --- | --- | C | GAGACATATT |
| SEQ 15 | GCAAGCAGCT | CTCTGTCGGT | GCCGTCGGCT | TGGTCAAC | --- | TCG | GCTGAGATCG | CCAAGGAGAC | CGTCCAGGAG | AAGGAGGATG | GCAGAGGTCAC |
| SEQ 17 | TCCCCAAGCT | TCCTCTCATG | GTACACCGCG | GCTTCCGC | --- | --- | --- | --- | --- | ACTC | GTCAGGTCAT |
| SEQ 18 | TCCCCAAGCT | TCCTCTCATG | GTACACCGCG | GCTTCCGC | --- | --- | --- | --- | --- | ACTC | GTCAGGTCAT |
| SEQ 20 | --- | AGGTT | GCTCATAGGC | GCGGTCGGCA | ACATCAAC | --- | --- | --- | --- | ACGG | CTGACATTGC |
| SEQ 21 | --- | AGGTT | GCTCATAGGC | GCGGTCGGCA | ACATCAAC | --- | --- | --- | --- | ACGG | CTGACATTGC |
| SEQ 23 | --- | AT | ACTACTTGGC | GCTGTCCGAA | TGATCAGG | --- | --- | --- | --- | GATG | GTCCTACGGC |
| SEQ 25 | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| SEQ 26 | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| SEQ 28 | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| SEQ 29 | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| SEQ 32 | --- | AGTGT | ACTTGTTTCA | GCAGTAGGTG | GAATCAAG | --- | --- | --- | --- | A | CTGGACATCT |
| SEQ 34 | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| SEQ 36 | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| SEQ 37 | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| SEQ 39 | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| SEQ 41 | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| SEQ 43 | --- | ATGGT | GGTCTACACC | ACCGCGCGCT | TCAAGACG | --- | --- | --- | --- | GTGGGGC | CCATGGTTCGA |
| SEQ 82 | --- | AAGAT | GTTGATCAGC | ACTGTTGGTA | GCATCAAG | --- | --- | --- | --- | ATAG | GTACCCCTTGC |
| SEQ 84 | --- | ATCGAACC | CGACGCGTGC | AAACGCATGC | TGCTCGGGG | --- | --- | --- | --- | CCGTGG | GAATGATGGA |

| | 1901 | 1911 | 1921 | 1931 | 1941 | 1951 | 1961 | 1971 | 1981 | 1991 |
|--------|-------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| SEQ 1 | GAATCAG--- | ---ATTCTAG | AGGAGCAG-- | | | | | | | |
| SEQ 2 | GAATCAG--- | ---ATTCTAG | AGGAGCAG-- | | | | | | | |
| SEQ 4 | CAATTCC--- | ---TTGTTGG | AGAAGGAC-- | | | | | | | |
| SEQ 5 | CAATTCC--- | ---TTGTTGG | AGAAGGAC-- | | | | | | | |
| SEQ 7 | CGAGGCAATG | CTGTCGGGAC | CTGAACCC-- | | | | | | | |
| SEQ 9 | GAACAAG--- | ---CTGCTTG | AGGAGGAG-- | | | | | | | |
| SEQ 11 | ATTGCTCAAC | AAATATTTAG | AAGAAGGA-- | | | | | | | |
| SEQ 13 | TAAACTCGAT | GAGTTTATTG | CTAATGGT-- | | | | | | | |
| SEQ 15 | CATCCAGCGC | GAGAACGGCG | CCAAGACT-- | | | | | | | |
| SEQ 17 | GGAGGCC--- | ---GCTTTGG | AATCCGAT-- | | | | | | | |
| SEQ 18 | GGAGGCC--- | ---GCTTTGG | AATCCGAT-- | | | | | | | |
| SEQ 20 | GC GCGATGTC | GTGGATGAGC | AGGGCGCCGA | GAAGGTGGCC | GAGGCCAAGC | AGACGCATGA | CACCATCGAG | GTCGTGAGCG | AATCATATGG | CGGCAAGACC |
| SEQ 21 | GC GCGATGTC | GTGGATGAGC | AGGGCGCCGA | GAAGGTGGCC | GAGGCCAAGC | AGACGCATGA | CACCATCGAG | GTCGTGAGCG | AATCATATGG | CGGCAAGACC |
| SEQ 23 | GAATGAAATC | CTAGAAAGTG | AAAAGCT-- | | | | | | | |
| SEQ 25 | | | | | | | | | | |
| SEQ 26 | | | | | | | | | | |
| SEQ 28 | | | | | | | | | | |
| SEQ 29 | | | | | | | | | | |
| SEQ 32 | TGCTGAA--- | ---GAGGTTT | TGCAATCT-- | | | | | | | |
| SEQ 34 | | | | | | | | | | |
| SEQ 36 | | | | | | | | | | |
| SEQ 37 | | | | | | | | | | |
| SEQ 39 | | | | | | | | | | |
| SEQ 41 | | | | | | | | | | |
| SEQ 43 | CGCGCTGCAG | GGCGTCGATG | GG--- | | | | | | | |
| SEQ 82 | GGAGGAG--- | ---ATCATCG | CTGGAAGAGA | GGACGATACC | | | | | | |
| SEQ 84 | AGGTTCC--- | ---TAGGATT | CGCCCAAC-- | | | | | | | |

| | 2001 | 2011 | 2021 | 2031 | 2041 | 2051 | 2061 | 2071 | 2081 | 2091 |
|--------|-------------|------------|------------|-------------|-------------|-------------|-------------|-------------|------------|-----------|
| SEQ 1 | GATATCGACG | TTGCGCTGGT | TGGCCGTGGG | TTCCAGAAGG | ATCCCGGTCT | GGCCTGGACG | TTTGCTCAGC | ACCTCGGCGT | C | |
| SEQ 2 | GATATCGACG | TTGCGCTGGT | TGGCCGTGGG | TTCCAGAAGG | ATCCCGGTCT | GGCCTGGACG | TTTGCTCAGC | ACCTCGGCGT | C | |
| SEQ 4 | GGACTGGACC | TTGTGCTGGT | TGGACGTGGC | TTCCAGAAGA | ACCCGGGGCT | GGTGTGGGCG | TGGGCCGACG | AGCTGAATGT | A | |
| SEQ 5 | GGACTGGACC | TTGTGCTGGT | TGGACGTGGC | TTCCAGAAGA | ACCCGGGGCT | GGTGTGGGCG | TGGGCCGACG | AGCTGAATGT | A | |
| SEQ 7 | AAGGCGGATG | CAATTCTGAT | AGCCCGTCAG | TTCTTGC CGG | AGCCAGAAATG | GGTGTTTTCC | ACGCGGAGAA | AGTTGGGCGT | G | |
| SEQ 9 | GGATTGGATG | TTGCGCTTGT | GGGACGTGGT | TTCCAGAAGG | ATCCCGGTCT | GGCGTGGACT | TTCCGCGCAGC | ATCTTGATGT | T | |
| SEQ 11 | ACATTTTGATC | TTGCTTTGAT | CGGTAGAGGA | TTTTTAAGAA | ATCCAGGTTT | GGTATGGGAG | TTTGCCGATA | AACTTGTGTT | T | |
| SEQ 13 | GACTTTTGATA | TGTCATTGAT | AGGTAAGGA | TTTCTCAAAA | ACACTGGATT | GATCAGCCGT | ATTGCTGACC | AATTGCAAGC | A | |
| SEQ 15 | CGTGGCGATA | TGGTCTTGT | TCCGACGCCG | TTCTTGAAGG | AGCCCGAGTT | CGTCTCTACT | GTGCGCGACG | AGTTGGGTGT | T | |
| SEQ 17 | GATTCGCAGA | TGATCGGTAT | CGGACGCCCG | GCCATCATCA | ACCTTTCTGT | TCCCGCCAACT | TTGATCTCTCA | ACCCGGAGGT | G | |
| SEQ 18 | GATTCGCAGA | TGATCGGTAT | CGGACGCCCG | GCCATCATCA | ACCTTTCTGT | TCCCGCCAACT | TTGATCTCTCA | ACCCGGAGGT | G | |
| SEQ 20 | AAGGCGGATC | TGGTCTCTAT | TGCTCGCCAG | TTCTTGC CGG | AGCCTGAGTT | TGTGTCTGAGG | ACGGCGCATC | ACCTTGGGGT | C | |
| SEQ 21 | AAGGCGGATC | TGGTCTCTAT | TGCTCGCCAG | TTCTTGC CGG | AGCCTGAGTT | TGTGTCTGAGG | ACGGCGCATC | ACCTTGGGGT | C | |
| SEQ 23 | ---GATG | TTACTTTTGT | CGCAAGGGAG | TTCTTAAGGA | ACCCGTCTGT | GGTGTCTAGAC | AGCGCGAACC | AGTTGGGTGA | A | |
| SEQ 25 | | | | | | | | | | |
| SEQ 26 | | | | | | | | | | |
| SEQ 28 | | | | | | | | | | |
| SEQ 29 | | | | | | | | | | |
| SEQ 32 | GGTATCGACA | TTGTGAGGGC | TGGACGTTTG | TTCCAAACGA | ATCTTGGTCT | GGTTCGAGCT | TTTGCTAACG | AGCTTGGCGT | G | |
| SEQ 34 | | | | | | | | | | |
| SEQ 36 | | | | | | | | | | |
| SEQ 37 | | | | | | | | | | |
| SEQ 39 | | | | | | | | | | |
| SEQ 41 | | | | | | | | | | |
| SEQ 43 | ---ATAGGCAT | CGGGCGCGCA | GCCGGTTCGG | AGCCGGACCT | CGCCAAAGAC | ATCATCGCGG | GCAAGGTGTC | CAGCATTTATC | AAATACGCCA | |
| SEQ 82 | CCCTTGGATC | TTGTGGCTTC | AGGCCGCTCT | TTCCAGAAGA | ACACTGGACT | TGTTTGGTCA | TGGGCTGACG | ATCTGAACAT | T | |
| SEQ 84 | GGCCACAGAC | CAGGCGAGAT | TGCGAAGTTG | GCCGACGAGT | CGATTACAGAG | CGGAGAGTGT | GATGCGGTAC | TTTGGCCACG | T | ---GGATTG |

| | 2401 | 2411 | 2421 | 2431 | 2441 | 2451 | 2461 | 2471 | 2481 | 2491 |
|--------|------------|------------|------------|------------|------------|-------------|------------|------------|------------|------------|
| SEQ 1 | CGTCCTCTTA | AGTTTCTCCG | TCATTGCTTC | TATTCTACTC | CAATCGCAAC | GCAATGGCGAC | CACGGATCGA | GTCGAATTTC | TCCGTCGTTT | GTATCTGATC |
| SEQ 2 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 4 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 5 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 7 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 9 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 11 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 13 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 15 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 17 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 18 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 20 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 21 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 23 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 25 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 26 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 28 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 29 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 32 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 34 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 36 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 37 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 39 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 41 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 43 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 82 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 84 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| | | | | | | | | | | |
| | 2501 | 2511 | 2521 | 2531 | 2541 | 2551 | 2561 | 2571 | 2581 | 2591 |
| SEQ 1 | AATATAAAAA | GCGGGGAATG | GCTTGACCCC | GCGCAGAATG | TCGATCTCTT | CGCAAACCTCT | CGGTGTATAG | GACGCTCAGC | AACGATCAAG | G |
| SEQ 2 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 4 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 5 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 7 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 9 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 11 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 13 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 15 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 17 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 18 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 20 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 21 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 23 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 25 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 26 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 28 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 29 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 32 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 34 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 36 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 37 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 39 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 41 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 43 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 82 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |
| SEQ 84 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |

Figure 2. A multiple alignments of the 2031 OR nucleic acid sequence from *A. fumigatus* (SEQ 1,2) along with related 2031 ORs from other fungi and bacteria (see also Example 4). Regions 1-11, marked with * or #, refer to regions conserved at the amino acid level between Ors but not OYEs.

Fungal 2031 ORs are given by SEQ ID No.: SEQ ID Nos. 1, 2, 4, 5, and 7, *A. fumigatus*; SEQ ID No. 9, *A. nidulans*; SEQ ID Nos. 11 and 13, *C. albicans*; SEQ ID Nos. 15, 17 and 18, *N. crassa*; SEQ ID Nos. 20, 21 and 43, *M. grisea*; SEQ ID No. 23 (NP_595868), *S. pombe*; SEQ ID Nos. 25 and 26, *C. trifolii*; SEQ ID Nos. 28, 29, 31, 32 and 34, *F. sporotrichioides*; SEQ ID Nos. 36, 37 and 82, *F. graminearum*; SEQ ID Nos. 39 and 41, *M. graminicola*; SEQ ID No. 84, *U. maydis*.

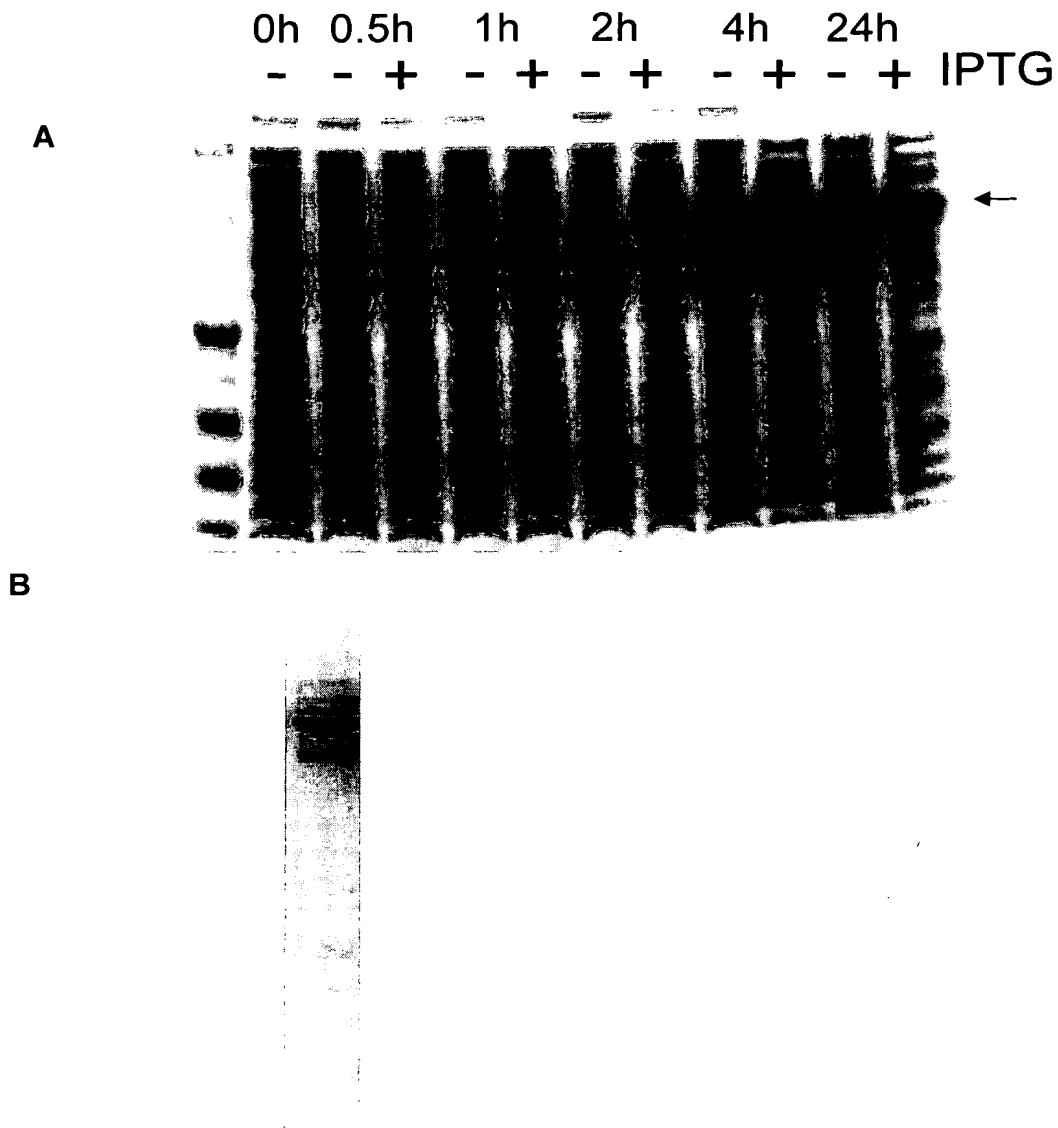


Figure 3. Recombinant 2031 OR. (A) Time course of recombinant 2031 OR induction over 24 hours after the addition of IPTG (samples without IPTG are also shown). The gel was stained with coomassie; A prominent band of the correct molecular weight (marked with an arrow) is seen. (B) Coomassie stained gel showing purified recombinant 2031.

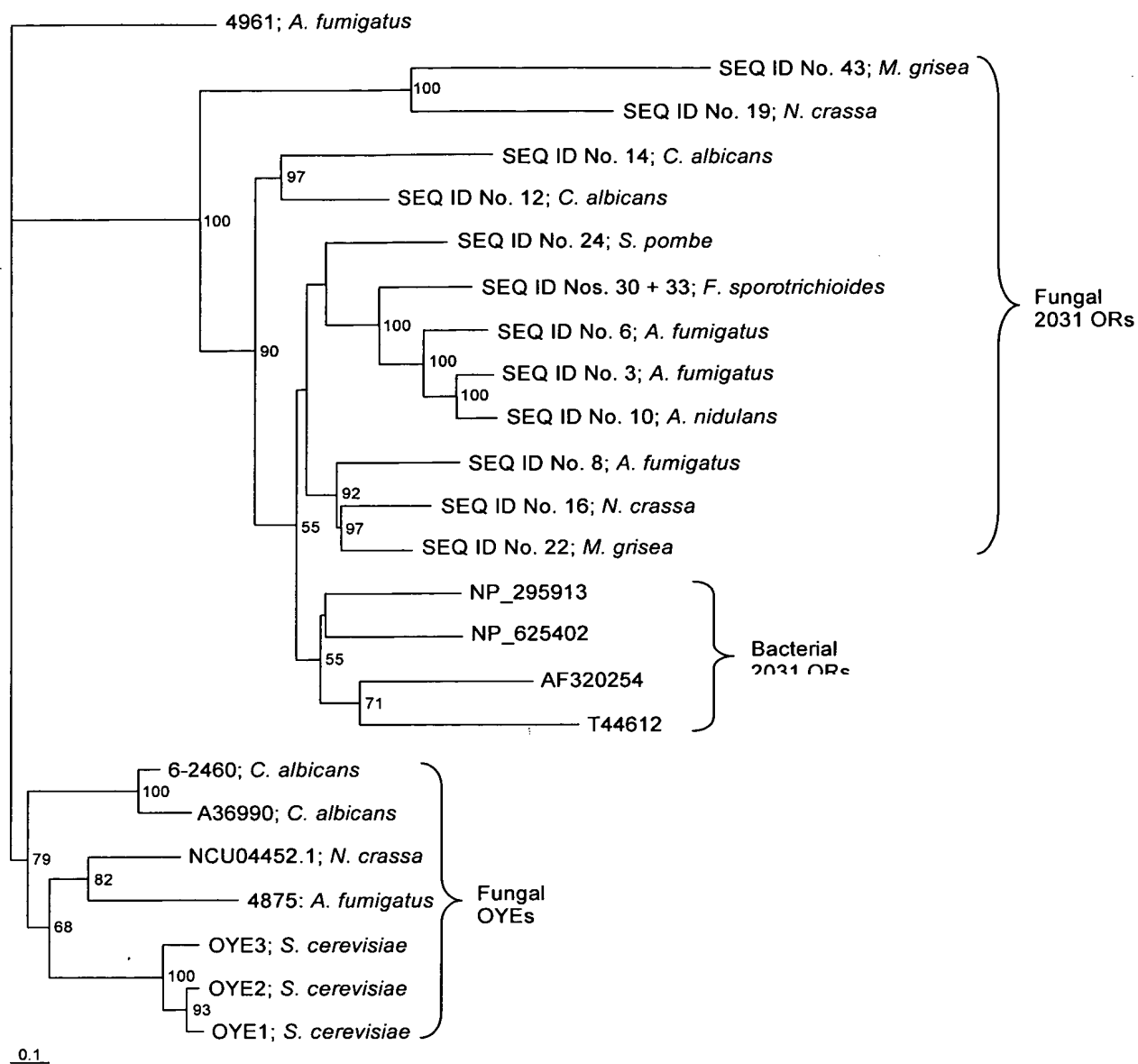


Figure 4. Phylogenetic tree showing relationships between *A. fumigatus* 2031 OR and similar proteins. This demonstrates a 2031 OR clade, which can be distinguished from the OYE proteins.

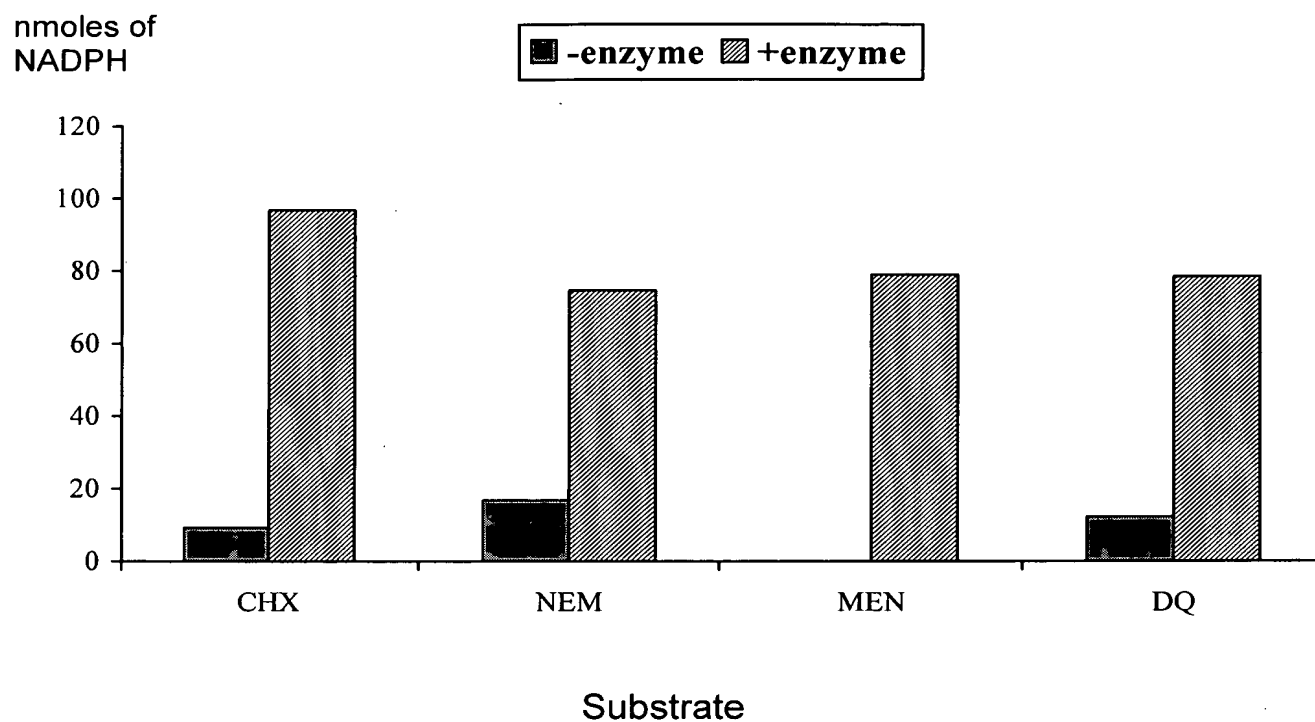


Figure 5: NADPH dehydrogenase activity of recombinant 2031 OR with cyclohexenone (CHX), N-ethylmaleimide (NEM), menadione (MEN) or duroquinone (DQ) as substrates. Final concentrations in the assay were as follows: 500 μ M substrate, 120 μ M NADPH, 1 μ g/200 μ L 2031 OR.

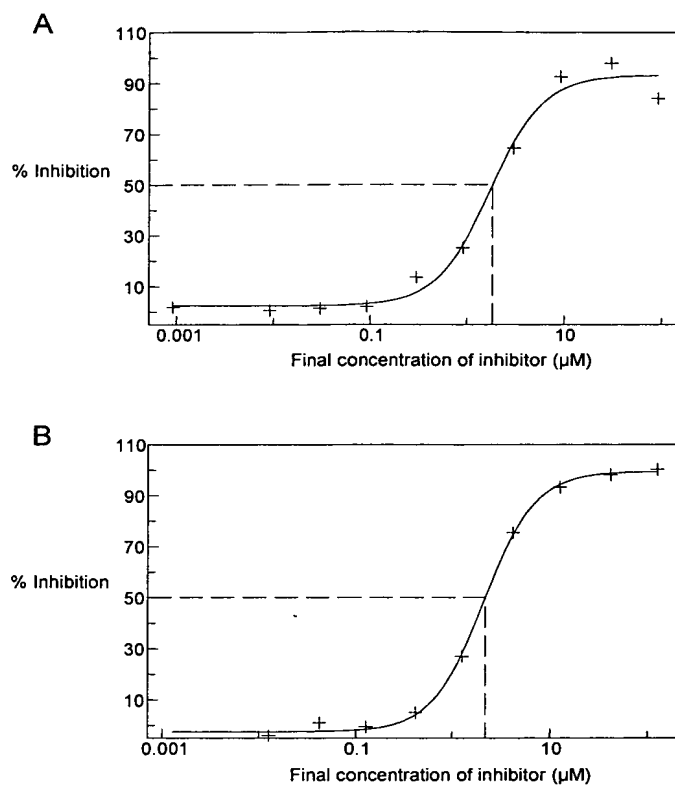


Figure 6: Inhibition of 2031 OR function by two inhibitors (shown in **A** and **B**) identified by high-throughput screening.